IMPERIAL INSTITUTE

OF

AGRICULTURAL RESEARCH, PUSA.

PROCEEDINGS

OF THE

ROYAL SOCIETY OF LONDON

SERIES A

CONTAINING PAPERS OF A MATHEMATICAL AND PHYSICAL CHARACTER.

VOL. CXLI.

LONDON:

PRINTED FOR THE ROYAL SOCIETY AND SOLD BY
HARRISON AND SONS, LTD., ST. MARTIN'S LANE,
PRINTERS IN ORDINARY TO HIS MAJESTY.

SEPTEMBER, 1933.

LONDON:
HARRISON AND SONS, LTD., PRINTERS IN ORDINARY TO HIS MAJESTY,
ST. MARTIN'S LANE.

CONTENTS.

SERIES A. VOL. CXLI.

No. A. 843.—July 3, 1933.

PAGE

The Heats of Combustion of Carbon Monoxide in Oxygen and of Nitrous Oxide in Carbon Monoxide at Constant Pressure. By J. H. Awbery and E. Griffiths, F.R.S. 1

A Bomb Calorimeter Determination of the Heats of Formation of Nitrous Oxide and Carbon Dioxide. By R. W. Fenning and F. T Cotton. Communicated by Sir Joseph Petavel, F.R.S 17

The Upper Pressure Limit in the Chain Reaction between Hydrogen and Oxygen. By G. H. Grant and C. N. Hinshelwood, F.R.S 29

The Thermal Decomposition of Acetaldehyde and the Existence of Different Activated States. By C. J. M. Fletcher and C. N. Hinshelwood, F.R.S 41

Notes on Some Electronic Properties of Conductors and Insulators. By R. H. Fowler, F.R.S. 56

The Relationship between Viscosity, Elasticity and Plastic Strength of a Soft Material as Illustrated by some Mechanical Properties of Flour Dough.—III. By R. K. Schofield and G. W. Scott Blair. Communicated by Sir John Russell, F.R S.... 72

The Behaviour of Electrolytes in Mixed Solvents. Part V.—The Free Energy of Lithium Chloride in Water-Alcohol Mixtures and the Salting-out of Alcohol. By J. A. V. Butler and D. W. Thomson. Communicated by J. Kendall, F.R S. 86

Contributions to the Mathematical Theory of Epidemics. III.—Further Studies of the Problem of Endemicity. By W. O. Kermack and A. G. McKendrick. Communicated by Sir Gilbert Walker, F.R.S 94

The Velocity of Sound in Gases in Tubes. By G. W. C. Kaye and G. G. Sherratt. Communicated by Sir Joseph Petavel, F.R.S 123

Conduction of Heat in Powders. By W. G. Kannuluik and L. H. Martin. Communicated by T. H. Laby, F.R.S 144

The Thermal and Electrical Conductivities of Several Metals between −183° C. and 100° C. By W. G. Kannuluik. Appendix by C. E. Eddy and T. H. Oddie. Communicated by T. H. Laby, F.R.S 159

The Deposition of Sputtered Films. By R. W. Ditchburn. Communicated by Sir Joseph Thomson, F.R.S 169

Determinations of the Signs of the Fourier Terms in Complete Crystal Structure Analysis. By K. Banerjee. Communicated by Sir William Bragg, O.M., F.R.S. 188

Collisions of α-particles with Fluorine Nuclei. By N. Feather. Communicated by Lord Rutherford, O.M., F.R.S. (Plates 1 and 2) 194

PAGE

Internal Photoelectric Absorption in Halide Crystals. By R. W. Gurney. Communicated by R. H. Fowler, F.R.S. 209

Boundary Conditions for the Wave Equation. By W. H. McCrea and R. A. Newing. Communicated by S. Chapman, F.R.S. 216

The Adsorption of Iodine by Potassium Iodide. By B. Whipp. Communicated by E. K. Rideal, F.R.S. 217

The Observation of Gravity by Means of Invariable Pendulums. By E. C. Bullard. Communicated by Sir Gerald P. Lenox-Conyngham, F.R.S. (Plate 3) 233

Experiments on the Transmutation of Elements by Protons. By M. L. E. Oliphant and Lord Rutherford, O.M., F.R.S. 259

No. A 844.—August 1, 1933.

Results of Calculations of Atomic Wave Functions. I.—Survey, and Self-consistent Fields for Cl^- and Cu^+. By D. R. Hartree, F.R.S. 282

Laboratory Determinations of the Magnetic Properties of Certain Igneous Rocks. By A. F. Hallimond and E. F. Herroun. Communicated by Sir Frank Smith, Sec. R.S. 302

New Methods of Deriving Stresses Graphically from Photo-Elastic Observations. By H. Neuber. Communicated by E. G. Coker, F.R.S. 314

On Approximation by Polygons in the Calculus of Variations. By L. C. Young. Communicated by G. H. Hardy, F.R.S. 325

The Infra-Red Absorption Spectrum of Nitrogen Tetroxide and the Structure of the Molecule. By G. B. B. M. Sutherland. Communicated by T. M. Lowry, F.R.S. 342

A New Presentation and Interpretation of the Quantum Equations. By H. T. Flint. Communicated by O. W. Richardson, F.R.S. 363

Experiments on the Protons produced in the Artificial Disintegration of the Nitrogen Nucleus. By E. C. Pollard. Communicated by R. Whiddington, F.R.S. 375

Ionization of Mercury Vapour by Positive Ions of Mercury and Potassium. By R. M. Chaudhri. Communicated by Lord Rutherford, O.M., F.R.S. 386

The Structure of Magnesium, Zinc and Aluminium Films. By G. I. Finch and A. G. Quarrell. Communicated by W. A. Bone, F.R.S. (Plates, 4, 5.) 398

The Catalytic Properties and Structure of Metal Films. Part I.—Sputtered Platinum. By G. I. Finch, C. A. Murison, N. Stuart and G. P. Thomson, F.R.S. (Plates 6, 7) 414

Free Paths and Transport Phenomena in Gases and the Quantum Theory of Collisions. I.—The Rigid Sphere Model. By H. S. W. Massey and C. B. O. Mohr. Communicated by P. A. M. Dirac, F.R.S. 434

Emission of Metallic Ions from Oxide Surfaces. I.—Identification of the Ions by Mobility Measurements. By L. Brata. Communicated by A. P. Chattock, F.R.S. 454

Emission of Metallic Ions from Oxide Surfaces. II.—Mechanism of the Emission. By C. F. Powell and L. Brata. Communicated by A. P. Chattock, F.R.S. 463

PAGE

The Scattering of Electrons by Metal Vapours. I.—Cadmium. By E. C. Childs and H. S. W. Massey. Communicated by Lord Rutherford, O.M., F.R.S. 473

The Influence of Pressure on the Spontaneous Ignition of Inflammable Gas-Air Mixtures. I.—Butane-Air Mixtures. By D. T. A. Townend and M. R. Mandlekar. Communicated by W. A. Bone, F.R.S. 484

No. A 845.—September 1, 1933.

Meeting for Discussion on the Ionosphere. Opened by E. V. Appleton, F.R.S. 697

An Optically Active Arsonic Acid Possessing Molecular Dissymmetry Resolution of dl-spirobis-3 : 5-*Dioxan*-4 : 4′-*di* (*phenyl-p-arsonic acid*). By C. S. Gibson, F.R.S. and B. Levin 494

Energy Relations in the β-Ray Type of Radioactive Disintegration. By C. D. Ellis, F.R.S. and N. F. Mott 502

Ocean Currents Produced by Evaporation and Precipitation. By G. R. Goldsbrough, F.R.S. 512

Probability and Chance in the Theory of Statistics. By M. S. Bartlett. Communicated by G. Udny Yule, F.R.S. 518

Experiments on the Raman Effect at very Low Temperatures. By G. B. B. M. Sutherland. Communicated by T. M. Lowry, F.R.S. 535

On the Electromagnetic Fields due to Variable Electric Charges and the Intensities of Spectrum Lines according to the Quantum Theory. By V. Fock. Communicated by P. A. M. Dirac, F.R.S. 550

The Coefficients of Absorption and Opacity of a Partially Degenerate Gas. By B. Swirles. Communicated by E. A. Milne, F.R.S 554

Adsorption, Oriented Overgrowth and Mixed Crystal Formation. By C. W. Bunn. Communicated by C. N. Hinshelwood, F.R.S. (Plate 8.) 567

X-Ray Analysis of the Crystal Structure of Durene. By J. M. Robertson. Communicated by Sir William Bragg, O.M., F.R.S. 594

The Spectra of the Halogen Molecules. Part I.—Iodine. By W. E. Curtis and S. F. Evans. Communicated by T. H. Havelock, F.R.S. (Plate 9.) 603

Experiments on Molecular Scattering in Gases. I.—The Method of Crossed Molecular Beams. By R. G. J. Fraser and L. F. Broadway. Communicated by T. M. Lowry, F.R.S. 626

Experiments on Molecular Scattering in Gases. II.—The Collision of Sodium and Potassium Atoms with Mercury. By L. F. Broadway. Communicated by T. M. Lowry, F.R.S. 634

The Emission of Electrons from Tungsten and Molydbenum under the Action of Soft X-Rays from Copper. By J. Bell. Communicated by R. H. Fowler, F.R.S. 641

The Flow Past Circular Cylinders at Low Speeds. By A. Thom. Communicated by G. I. Taylor, F.R.S. (Plates 10–13.) 651

PAGE

On the Ionization of Light Gases by X-Rays. I.—Technique. By W. R. Harper. Communicated by Lord Rutherford, O.M., F.R.S. 669

On the Ionization of Light Gases by X-Rays. II.—The Ionization of Hydrogen by Recoil Electrons. By W. R. Harper. Communicated by Lord Rutherford, O.M., F.R.S. 686

The Transmutation of Lithium by Protons and by Ions of the Heavy Isotope of Hydrogen. By M. L. E. Oliphant, B. B. Kinsey and Lord Rutherford, O.M., F.R.S. 722

A Photographic Investigation of the Transmutation of Lithium and Boron by Protons and of Lithium by Ions of the Heavy Isotope of Hydrogen. By P. I. Dee and E. T. S. Walton. Communicated by Lord Rutherford, O.M., F.R.S. (Plates 14–17.) 733

Index 743

Title, Contents, etc.

PROCEEDINGS OF
THE ROYAL SOCIETY.

Section A.—Mathematical and Physical Sciences.

The Heats of Combustion of Carbon Monoxide in Oxygen and of Nitrous Oxide in Carbon Monoxide at Constant Pressure.

By J. H. Awbery, B.A., B.Sc., F.Inst.P., and Ezer Griffiths, D.Sc., F.R.S., Physics Department, National Physical Laboratory, Teddington.

(Received February 3, 1933.)

Introduction.

Data on the heats of combustion of nitrous oxide and of oxygen in carbon monoxide are not only of thermo-chemical importance in themselves but they also afford a basis for evaluating the heat of formation of nitrous oxide from its elements.

The present paper is concerned with the measurement of these two heats of combustion by the constant pressure method. The reaction was brought about by the continuous burning of one of the reacting gases in an excess atmosphere of the other. The heat evolution was obtained by measuring the temperature rise of the water in a calorimeter, and the weight of gas involved in the reaction was determined by absorbing and weighing the carbon dioxide produced.

The Calorimetric System.

The Calorimeter Vessel.—The calorimeter was made of brass, with water as the calorimetric fluid, and it was surrounded by a water jacket so that it could be used adiabatically.* The arrangement is shown in fig. 1. The outer cylindrical jacket is hollow and is connected with the two halves of the lid by rubber tubing, so that the same jacketing water can be circulated round the whole of the outer system by means of the simple hand-worked pump† seen on the right of the drawing; a system of baffles in the jacket and lid compels the water to circulate completely round the interspace and leave no "dead water" regions. On the outer wall of the jacket a nichrome heating coil is wound directly on the metal, experience having shown that the oxide film on nichrome is sufficient electrical insulation to prevent short-circuiting of the coil. This is thermally insulated on the outside by asbestos paper maintained in place by strips of tape. By means of a switch the heat supply to the jacket can be stopped and started at will, so as to keep the jacket temperature oscillating slightly about the temperature of the inner calorimeter. As an indicator for this purpose, a multiple copper-constantan thermo-couple with three junctions in each limb is arranged with one limb in the water of the jacket and the other in that of the calorimeter itself. The thermo-couple is

---

* The calorimeter and jacket were constructed in the workshop of the Engineering Department of the Laboratory.

† A possible refinement would have been a double acting pump so as to maintain a continuous flow.

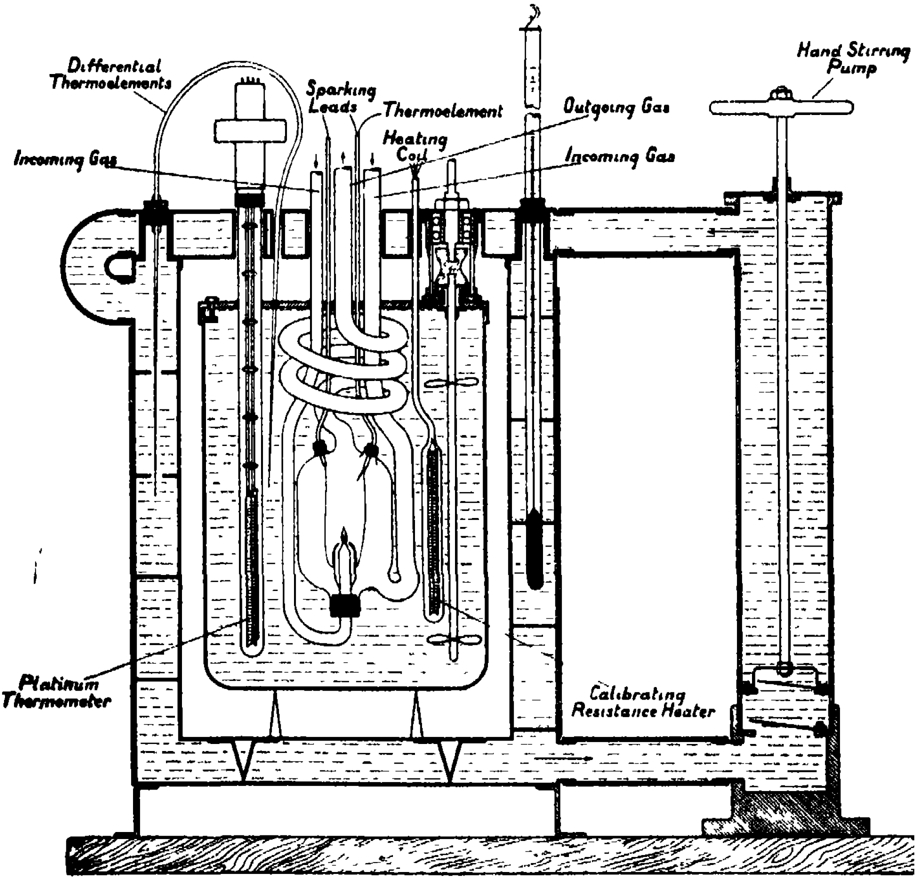


Fig. 1.

connected directly to a mirror galvanometer throwing a spot of light on a scale. It was found quite practicable to keep the two temperatures by hand control within 0·02° C. and generally within 0·01° C. during the half-hour required for an experiment. In addition to the thermo-couple which shows the difference between jacket and calorimeter, there is a calorimetric thermometer divided to 1/100° C. for showing the actual temperature of the water in the jacket.

The calorimeter itself is a cylindrical brass vessel 8 inches high and 5·5 inches in diameter, and holding about 2·6 litres of water; it was supported on three ebonite pegs inside the jacket. It was also found necessary to use three small pieces of cork for " centering " the calorimeter, since any slight lateral movement might have fractured the glass tubes which had of necessity to pass through the two lids. The lid of the calorimeter was made in one piece, with a slot cut out as seen in fig. 1. The joint between the calorimeter and the lid was a metal to metal cone with a thin film of grease to ensure good thermal contact between the lid and the cylindrical vessel. The glass tubes projecting through the lid were allowed to follow their natural directions without forcing, and then the aperture through the lid was sealed and made impervious to water vapour, by using a plastic cement. The calorimeter contained :—

(*a*) The reaction vessel in which the gases were burnt (described in the next section).

(*b*) A platinum resistance thermometer.

(*c*) A stirrer with two propellers on a hollow shaft rotated by means of an electric motor. Metallic connection between the propeller shaft and the motor shaft was broken by the interposition of an insulating coupling. The shaft had an " umbrella " dipping in oil which sealed the opening into the calorimeter.

(*d*) A heater wound with platinum wire on a mica frame, and enclosed in a silica envelope.

(*e*) One arm of a differential couple of which the other arm was in the jacket.

The Gas Circuits and Reaction Vessel.—The two gases were taken from cylinders—see fig. 2—and, after passing through soda-asbestos to remove carbon dioxide and through phosphorus pentoxide to remove water, each was led through a simple flowmeter, and then past mercury thermometers. The gas which formed the atmosphere around the flame proceeded direct to the reaction vessel, whilst the path for the other gas was selected by a two-way tap which could send it either (*a*) out through a long glass tube into a ventilating duct to be discharged clear of the building, or (*b*) into the jet, after sparking had been commenced. The connecting tubes whenever possible were of glass, but at certain points flexible tubes were necessary, and rubber was used. It was kept as short as possible, and rendered impervious by painting the tubing

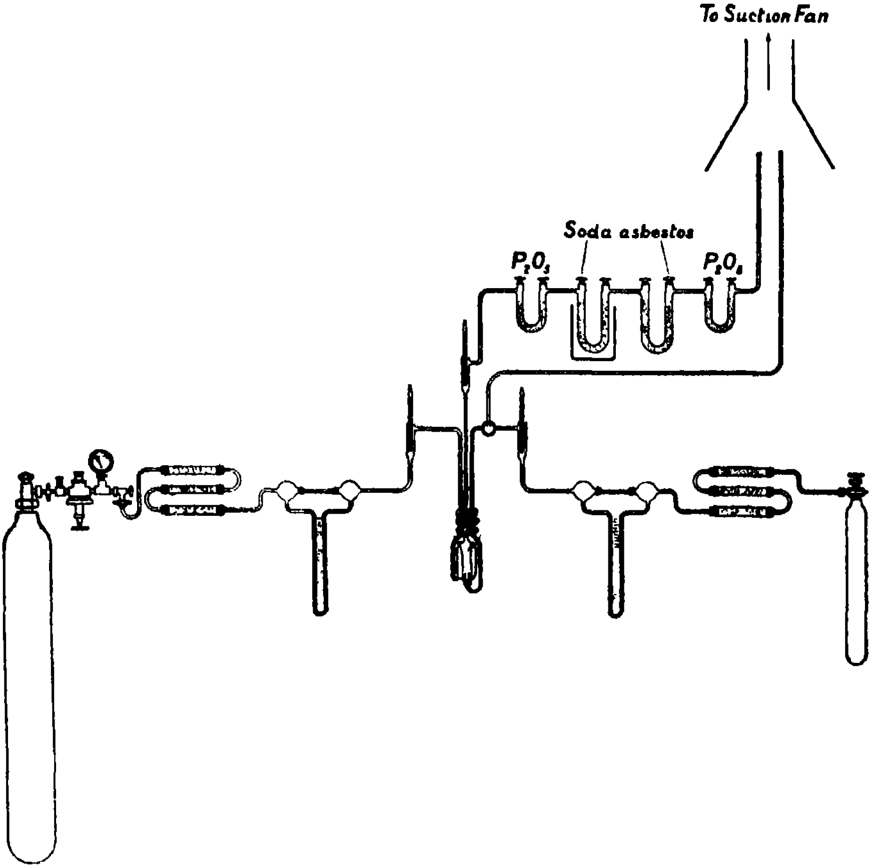


Fig. 2.

as well as the joint between the rubber and glass, with an alcoholic solution of sealing-wax. The reaction vessel consists essentially of a pyrex glass chamber about 3 inches high and 1·8 inches in diameter with a pyrex " burner " or " jet " for the flame. One gas was led to this burner and the other, which formed the atmosphere during the combustion, was led to the tube which opens into the base of the reaction vessel. The products of combustion passed out through the spiral of 0·3-inch bore tubing, which was wholly immersed in the water of the calorimeter, so as to ensure that the heat of the reaction was conveyed to the calorimetric fluid.

Combustion was started by sparking between two platinum wires as shown in the figure. A thermocouple was also introduced into the reaction vessel above the jet and slightly to one side. The millivoltmeter connected to this served to indicate when the gases were first ignited, and also gave warning when for any reason the flame happened to be extinguished in the course of an experiment. Much difficulty was encountered in the preliminary experiments in determining the conditions for a steady flame. It was found that the insertion of a plug of cotton-wool in the tube leading to the jet was essential for the maintenance of the flame.

After leaving the reaction vessel, the mixed products of combustion flowed past a mercury thermometer and then proceeded direct to the absorption system.

Absorption System for Carbon Dioxide.—This consisted of two U-tubes filled with soda-asbestos, and two of phosphorus pentoxide. The first soda-asbestos tube, in which the main part of the absorption took place, was immersed in a water bath to keep it cool. The second absorption tube increased in weight by about ½ per cent. of the first.

Heat Capacity of the Calorimeter and Contents.

The heat capacity was determined in terms of electrical units by dissipating energy inside the calorimeter, using the platinum heater already mentioned. The external resistance in the heater circuit was adjusted beforehand so that when the current was switched on to the heating coil the rate of production of energy should be approximately the same as that in the combustion experiments. The current was then switched on, and the time at which this occurred was automatically recorded on the tape of a chronograph, the circuit of the latter being completed through one of the knife-edges of the switch. On the same tape a time record was marked. Since the heater was of platinum, its resistance gradually rose as it warmed up, so that the experiment was divided into two sections, an initial period which lasted about 100 seconds, during which the heater resistance was changing, and the settled period lasting about 1600 seconds, during which the resistance of the coil was constant. During the first period the current flowing through the heating coil and the potential difference across its ends were observed on a moving coil instrument. The time of each reading was marked by a second pen on the chronograph tape, actuated by the observer pressing a tapping key as he noted an instrument reading. In the second period, current and potential were measured by means of a potentiometer, through the intermediary, respectively, of a standard 1·0 ohm shunt and a volt box giving a reduction of 200 to 1. At the termination of the experiment, the opening of the switch of the heating circuit was recorded on the chronograph tape.

Corrections were applied for the heating due to the stirrer in the calorimeter, and for the heat generated in the current-carrying leads of the heating coil. A further correction on the energy measurements was required since both the volt box of the moving coil instrument and the volt box for the measurements by potentiometer, were in parallel with the heater coil, and consequently the measured current was greater than that actually flowing through the heater. The corrections were calculated from the known resistance of the volt boxes on the ranges used, together with the resistance of the heater. The correction to the energy supply on this account was of the order of 1 part in 400 or 500. The correction for the heat developed in the current leads had about twice the magnitude of the above-mentioned correction. It was evaluated from a knowledge of the specific resistance of the wire used as leads, but it must be noted that there is a slight uncertainty in determining what is the correct length of wire which must be taken as functioning usefully in this connection. There is a space between the lid of the calorimeter and the lid of the jacket, and clearly the heat developed in this part of the lead is divided, some ultimately finding its way to the calorimeter and some to the outside. Only the latter part should be applied as a correction to the observed energy. As an approximation, one-half of the wire in the space referred to was taken as effective. The aggregate effect of these various corrections together with possible random errors leaves the accuracy which can be claimed as perhaps 1 part in 500 for each individual experiment.

It might be noted that there would be some advantage in using leads of low resistance jacketed along their length with water at the temperature of the calorimeter enclosure to minimize heat conduction from the calorimeter. This, however, would introduce considerable constructional complications.

In the above described determinations of the electrical equivalent, conditions differed as little as possible from those prevailing during a combustion experiment, except that in the latter there was of necessity a current of gas flowing through the reaction vessel. A special calibration experiment was carried out to ascertain whether the gas flow did affect the water equivalent after a correction for the heat carried off by the exit gas had been applied in the way described in the section on combustion experiments. In this experiment, carbon monoxide was passed through the apparatus whilst the calibration was carried out. The result agreed with the "no gas" calibration within the limits of the possible accuracy of the experiments.

Other checks on the satisfactory behaviour of the apparatus were provided by carrying out experiments for varying periods of time with approximately the same rate of energy input, and also experiments in which the rate of supply of energy was varied by alteration of the current through the heater. Some of these experiments are quoted in Table I. The quantity of water in the calorimeter had been changed between the dates concerned, and is also different from that in the combustion experiments.

Table I.

Date.	Duration of experiment (seconds)	Temperature rise. (° C.)	Mean temperature (° C.)	Energy rate. (watts)	Electrical equivalent (Joule/1° C.)	Conditions varied.
13.11.31	2523	2·32	22·61	5·5	11810	Energy rate.
5.11.31	1381	5·07	22·61	21·7	11816	
4.1.32	2314	6·28	22·41	16·0	11732	Duration of experiment
5.1.32	2404	6·30	22·26	15·4	11725	
7.1.32	1652	4·21	21·54	15·1	11721	

Electrical calibrations were generally, but not invariably, alternated with combustion experiments, but the result of each combustion experiment was calculated, using the mean of all the calibrations performed during the group of experiments to which it belonged.

The mass of the calorimeter and metal and glass parts was 2513 gm. which would increase in electrical equivalent by about 0·5 joules per degree Centigrade for each degree rise in temperature. The thermal capacity of the water would decrease by about 1·5 joules per 1° C. for a degree rise in the neighbourhood of 25° C. so that, on the whole, the correction for temperature is —1 joule per degree Centigrade for a rise of 1° C.

Combustion Experiments.

Carbon Monoxide and Oxygen.—The general plan during the combustion experiments has already been outlined. One of the gases, after being dried, purified from carbon dioxide, and metered, was led to the burner in the reaction vessel, whilst the other gas, in considerable excess of the theoretical quantity required for combustion of the first gas, was also purified and metered, and then passed into the reaction vessel to provide an atmosphere for the combustion. The products of combustion, together with the excess of the second gas, passed out through the pyrex glass spiral and into the absorption tubes. Temperatures of the inlet and exit gases were taken by means of mercury thermometers.

The procedure during a combustion experiment was to allow the gas which was to be in excess to run for a few minutes to flush out the system; during this period, the other gas was directed, by means of the two-way tap referred to previously, so as to pass to waste without entering the combustion chamber. The spark was then started and the two-way tap turned so as to pass the gas to be burnt into the burner of the reaction vessel. A stop-watch was started as the spark was switched on and allowed to run until the millivoltmeter connected to the thermo-element gave a deflection indicating that the gases were burning, when the spark was switched off (usually after 10 seconds). Readings of the flowmeters, room temperature, times and gas temperatures (entrance and exit) were taken periodically throughout the experiment. At the conclusion of the experiment, the flame was extinguished by turning the two-way cock so that it stopped the supply of gas to the jet. The gas forming the atmosphere was allowed to flow for about a quarter of an hour so as to wash out all carbon dioxide into the absorption tubes, and these were then closed and allowed to come into temperature equilibrium with the room prior to weighing

Whenever the absorption tubes were weighed, a dummy absorption tube was placed on the other pan of the balance, to compensate for variations in the moisture absorbed on the surface of the glass, and also for variations in the buoyancy of the air.

In the reaction between oxygen and carbon monoxide, each gas in turn was burnt in an atmosphere of the other, and the agreement of the results is a check on the general accuracy of the method.

The carbon monoxide employed was prepared by the Chemical Research Laboratory of the Department of Scientific and Industrial Research and analyses were furnished with it. The impurities present were hydrogen* and nitrogen, the quantities being of the order of 0·03 per cent. hydrogen and

---

* This was determined by oxidation to water, using copper oxide. Its presence is probably essential to the combustion of the gases.

0·7 per cent. nitrogen. The oxygen was stated by the suppliers to be of atmospheric origin, so that its impurities would be nitrogen, together with argon, etc., and possibly water vapour and carbon dioxide. The carbon dioxide and water were removed from the gases by absorption; that this absorption was complete was checked by passing the mixed gases through the whole apparatus, but without igniting them. No increase in weight could be detected in the absorption tubes at the exit, provided the apparatus had not been left standing for some time.

Nitrogen is an inert gas, and its presence could cause no error in the results, since it would affect neither the weight of the absorbed products, nor the heat generated in the combustion, unless it reacts with the oxygen to form one of the oxides of nitrogen. The possible formation of nitrogen oxides was tested for, by means of a modification of the "brown ring" test and none could be detected.

Thus the only impurity requiring special attention is hydrogen. This would react with the oxygen, forming water with the evolution of heat and, if the products were passed directly into soda-asbestos, the water would be absorbed, so increasing the apparent amount of carbon dioxide absorbed, and therefore leading to a wrong estimate of the carbon monoxide burnt. For this reason, a tube of phosphorus pentoxide was placed in the exit train, before the soda-asbestos tubes, so as to remove any water. A second phosphorus pentoxide tube was placed at the exit of the absorbing train, to prevent the escape of any water from the soda-asbestos tubes themselves. The latter tube usually increased in weight by 0·005 gm. or less, whilst the amount of carbon dioxide absorbed was of the order of 8 gm. The absorption of the small quantity of water in the last phosphorus pentoxide tube would thus seldom affect the result by 1 part in 1600 even if completely ignored. The water vapour collected in the first absorption tube cannot be considered as that formed solely by the combustion of the hydrogen present as impurity in the gases, since some of of it may have come out of the reaction vessel or the connecting tubes of the apparatus. It was assumed that the whole of the hydrogen present in the carbon monoxide consumed had reacted with oxygen to form water. A correction on this basis was made to the observed heat generated. No correction to the weight of carbon dioxide is necessary, since the water was removed before the products of combustion reached the soda-asbestos tubes.

The details of a typical experiment will best illustrate the order of magnitude of the various quantities involved.

Experiment No. 2 *of May* 18, 1932, *Carbon Monoxide Burnt in Oxygen.*

	gm.
Gain in weight of soda-asbestos tube No. 1	$8\cdot150_5$
,, ,, ,, No. 2	$0\cdot003$
,, ,, P_2O_5 tube at exit	$0\cdot002$
Total	$8\cdot155_5$

From the equation $2CO + O_2 = 2CO_2$, it follows that 28·00 gm. of carbon monoxide burn to form 44·00 gm. of the dioxide Hence,

$$CO \text{ burnt} = \frac{8\cdot1555 \times 28}{44} = 5\cdot2 \text{ gm.}$$

The rate of flow of oxygen (obtained from a calibration of the flowmeter) was 1·77 cubic feet per hour = 1·77 × 28·3 litres per hour. The duration of the experiment was 27 minutes, so that the oxygen used was

$$1\cdot77 \times 27 \times 28\cdot3/60 = 22\cdot5_5 \text{ litres.}$$

Taking this to be at room temperature (20·4° C.) and taking 1 molecular weight to occupy 22·4 litres at N.T.P. this corresponds to a weight of

$$\frac{22\cdot5_5 \times 273\cdot1 \times 32}{22\cdot4 \times 293\cdot5} = 30\cdot0 \text{ gm. oxygen.}$$

Adding to this the weight of carbon monoxide gives the mass of gas passed through the apparatus as 35·2 gm. with a mean specific heat 0·222 calories per gram per 1° C.

From the mercury thermometers in the gas lines, the mean temperature of the carbon monoxide at inlet was $20\cdot7_8$° C. and that of the oxygen $20\cdot5_9$° C. The exit gases had a mean temperature of $26\cdot6_1$° C. Thus in passing through the apparatus, the carbon monoxide was warmed* $5\cdot8_3$° C. and the oxygen by $6\cdot0_2$° C. Weighting these in proportion to the quantity flowing, we may say that 35·2 gm. of gas were heated through 5·99° C. and hence must have abstracted 5·99 × 35·2 × 0·222 × 4·185 joules from the only source available—the calorimetric system. Hence this correction is 196 joules.

The temperature of the calorimeter was $24\cdot467_5$° C. initially and $28\cdot760_9$° C. finally, giving a rise of $4\cdot293_4$° C. of which $0\cdot012_6$° C. is due to the spark and $0\cdot027_2$° C. to the stirrer. The electrical equivalent of the calorimeter at the

* It might be feasible to reduce this correction very materially if the gases were circulated through a spiral immersed in the jacket water before entering the calorimeter.

mean temperature of 26·6° C. was 12263 joules per 1° C. hence the total energy produced by the combustion was $(12263 \times 4\cdot253_6) + 196 = 52358$ joules.

In this particular case, the carbon monoxide used contained 0·03 per cent. by volume of hydrogen. The correction for this was evaluated by assuming that all the hydrogen associated with the 5·2 gm. of carbon monoxide used was converted to water and that the water formed remained in the gaseous state. Thus 5·2 gm. CO = 5·2/28 moles CO contained $0\cdot03 \times 5\cdot2/2800$ moles H_2; but the (lower) heat of combustion of hydrogen is 281 kilo joules per mole, so that the heat produced in burning the hydrogen was $281 \times 0\cdot03 \times 5\cdot2/2800 = 16$ joules.

Thus the heat actually produced by the reaction between carbon monoxide and oxygen was 52358 − 16 = 52342 joules, and this produced $8\cdot155_5$ gm. of carbon dioxide. Consequently the heat of reaction for 1 mole of CO_2 is $52342 \times 44\cdot00/8\cdot155_5 = 282390$ joules, at a mean temperature of 26·6° C. The heat evolution increases 7 joules per mole for each 1° C. rise in temperature. Hence the heat of reaction corrected to 20° C. is 282390 − 50 = 282340 international joules per mole.

Results.

All four experiments in which carbon monoxide was burnt in oxygen are summarized in Table II.

Table II.

Date	Hydrogen content of CO.	Mean temperature	Heat of reaction corrected to 20° C.
	per cent.	° C	
27.4.32	0·03	21·8	282340
2.5.32	0·03	22·3	283100
18.5.32	0·03	23·1	283170
18.5.32	0·03	26·6	282340

We adopt a weighted mean of 282750 international joules per mole.

The deviations of the four experiments from this value, in parts per 1000, are 1·4, 1·2, 1·5 and 1·4 respectively.

As a check on the various corrections, and on the possibility of side-reactions having been overlooked, a few experiments were carried out with oxygen burning in carbon monoxide. The method of calculation was the same as that previously adopted, and the results are summarized in Table III. In giving weight to the results, an experiment was assessed as being less accurate if the gas failed to ignite as rapidly as usual, so that the correction for the heat generated in the sparking was greater than normal.

Table III.

Date	Hydrogen content of CO.	Mean temperature.	Heat of reaction corrected to 20° C.
	per cent.	° C.	
5 5.32	0·03	21·4	282560
6 5.32	0·03	20·5	282010
20 5.32	0·00	24·2	283290

Our weighted mean from this series is 282630 international joules per mole, the deviations, in parts per 1000, from this mean being 0·2, 2·2 and 2·3 respectively. When combining the two series, the second is entitled to less weight, since it contains fewer experiments, and also because the deviations are larger.

The correction for the hydrogen burnt is probably less certain when the hydrogen exists in the gas used as an atmosphere, than when it is in the gas burnt at the jet. We assign therefore a weight of 4 to the first series, and 1 to the second, hence obtaining the *final value* : 282730 *international joules per mole*, which is equivalent to $67\cdot5_7$ kg.—calories 15° per mole.

From a consideration of the various component possibilities of error it does not appear that the combustion experiments should be in error among themselves by more than 1 part in 600 ; this is in accord with the deviations from the mean found in the actual experiments quoted in the tables. The maximum possible error in these results, is therefore the sum of those in the calibrations and in the combustions, say, 3 parts in 1000. This takes no account of the fact that each series contains more than one experiment, so that the random errors would tend to cancel in the means, and the ultimate accuracy would be higher. It may be significant that the means of the two series agree to 1 part in 2000.

Comparison with other Observers.

Two recent determinations of this quantity have already been published. (1) Rossini* obtains a value which, reduced to 20° C. is 282960 joules per mole. This value is in excellent agreement with the value 282730 given above, the difference being only 0·8 parts per 1000. (2) Roth and Banse† give 67·86 cals. at 20° C. This differs from our value by nearly 5 parts in 1000.

---

* 'Bur. Stand. J. Res.,' vol. 6, p. 37 (1931).

† Landolt-Bornstein's 'Tables,' 2nd Erganzungs-band.

In a paper which is being published simultaneously with the present one, Fenning and Cotton (p. 17) give the value $67 \cdot 65_5$ cals. which is equivalent to 283090 joules. It was determined by the explosion method and is relative to an assumed value for the heat of formation of water. The divergence between the present value of 282730 and that of Fenning and Cotton is about 1 part in 800.

Nitrous Oxide and Carbon Monoxide.

The general plan of these experiments was identical with that of the oxygen, carbon monoxide experiments, but it was not found practicable to burn carbon monoxide in an atmosphere of nitrous oxide. In other respects, however, conditions were varied to a greater extent than in the experiments on carbon monoxide and oxygen. No purpose would be served by setting out a typical experiment, the calculations being of exactly the same type as before. The variation with temperature in the heat of reaction is much smaller, the increment being only 1·4 joules per mole for a rise of 1° C. In the first two experiments, the nitrous oxide used was of high "commercial" purity, and in the remainder it was exceptionally pure. Both samples were supplied by Messrs. Coxeter and the analysis of the gaseous phase of the purer material as given by them is set out below :—

N_2O.

	per cent.
Nitrogen Oxygen } less than	0·01
CO	absent
CO_2, less than	0·005
Water vapour, less than	0·005
N_2O, more than	99·98

The results are set out in Table IV and will show the variations in mean temperature, hydrogen content, correction for the heat carried off by the gases, etc. Further, the first two experiments were carried out with a different amount of water in the calorimeter from the later experiments and the nitrous oxide was a different sample.

The mean of these results is 364420, but we adopt the value 364340 joules per mole obtained by assignment of half weight to those experiments marked for various reasons as "less accurate."

The estimate of reliability given for the carbon monoxide, oxygen experiments should apply here, save that the correction for energy carried off by the

Table IV.

Date.	H_2 content of CO (per cent.)	Mean temperature (° C.)	Energy carried off by gases (Joules)	Temperature rise of calorimeter. (° C.)	CO_2 produced. (gm.)	Heat of reaction corrected to 20° C. (Joules/mole.)	Remarks.
26.5.32	0·06	22·2	169	3·75	5·56	365740	Less accurate.
26.5.32	0·06	26·0	456	5·04	7·47	366160	Less accurate.
30.5.32	0·06	20·3	101	4·12	6·06	364490	Calorimeter refilled. Another sample of N_2O
30.5.32	0·06	24·4	247	4·18	6·17	364080	—
6.6.32	0·06	19·1	82	4·56	6·74	362530	Less accurate
7.6.32	0·06	20·1	77	4·36	6·42	363910	—
14.6.32	0·044	24·0	111	3·91	5·76	364300	Another sample of CO.

gases is generally larger. We may probably say that it introduces no error larger than 1 part in 1000.

It is to be noticed, moreover, that since all the results with nitrous oxide were obtained by burning nitrous oxide in carbon monoxide, the correction for hydrogen is perhaps less certain; it is comparable, however, with the series of experiments when oxygen was burnt in carbon monoxide, and that series gave results in agreement with the other series where carbon monoxide was burnt in oxygen.

The deviations of the individual experiments from the adopted value in parts per 1000 are 3·1, 5·0, 0·4, 0·7, 5·0, 1·2 and 0·1.

Remembering the considerable variations in conditions, these deviations would leave little doubt that the mean itself is reliable to 2 parts or 3 parts in 1000, and this is the accuracy which is claimed for the work.

Comparison with other Observers.

The only recent determination known is one by Fenning and Cotton (p. 17). The value they found was 87·40 kg.-calories per mole, *i.e.*, 365715 joules per mole. The difference from our value is nearly 3 parts in 1000.

The Heat of Formation of Nitrous Oxide.

It is possible from this investigation to determine the heat of formation of nitrous oxide. Collecting the data in the form of thermo-chemical equations, we obtain:

$$CO + \tfrac{1}{2}O_2 = CO_2 + 282730.$$

$$CO + N_2O = CO_2 + N_2 + 364340.$$

By subtraction,

$$N_2O = \tfrac{1}{2}O_2 + N_2 + 81610,$$

i.e., the heat of formation of 1 mole of N_2O is 81610 international joules, or taking 4·186 absolute joules = 1 calories (15° caloric)[*] and 1 international joule = 1·00039 absolute joules—the heat of formation of N_2O is 19·50 *kg.-calories per mole.*

Berthelot and Thomsen have both determined the heat of formation of nitrous oxide by the reactions adopted here; Thomsen also deduced it by burning first oxygen and then nitrous oxide, with hydrogen. A determination by the latter method has also been published recently by T. C. Sutton. Messrs. Fenning and Cotton in the investigation previously referred to, determined the heat of formation of nitrous oxide by the explosion method, using first carbon monoxide and oxygen, and then carbon monoxide and nitrous oxide. They have also used hydrogen instead of carbon monoxide as the auxiliary gas, obtaining the same result. The various values are collected in Table V.

Table V.

Author.	Date.	Heat of formation of N_2O (Kg.-cals per mole).
Berthelot	1880	−20·6
Thomsen	1882	−17·7[*]
Sutton	1932	−20·5
Fenning and Cotton	1933	−19·74
Awbery and Griffiths	1933	−19·5

[*] The mean of values obtained using two different methods. One gave 18·0 and the other 17·5 kg.-cals per mole.

Acknowledgments.

Our thanks are due to various colleagues who have facilitated our work. Helpful advice was at all times freely given by members of the Staff of the Chemical Research Laboratory and we would mention in particular Mr. R. Taylor, M.A., B.Sc. We wish also to thank the Director of the National Physical Laboratory and the Superintendent of the Physics Department for their continued interest in the progress of the investigation. Mr. A. R. Challoner, Senior Observer in the Physics Department, assisted with the observational work and his manipulative skill was of great value in the assembly work.

---

[*] 'Nature,' vol. 128, p. 485 (1931).

Summary.

The heats of combustion of nitrous oxide and of oxygen in carbon monoxide have been determined by the constant pressure method. Each of these thermal quantities is obtained by burning one of the first-mentioned gases in an atmosphere of carbon monoxide, using an adiabatic calorimeter to measure the heat evolved. In addition, experiments were carried out in which carbon monoxide was burnt in an excess atmosphere of oxygen. The extent of the reactions was found by absorbing and weighing the carbon dioxide produced.

The heat capacity of the calorimeter and its contents was determined by generating a measured amount of energy in an electric heater in the calorimeter, and observing the corresponding temperature rise.

The result for the heat of combustion of carbon monoxide in oxygen was found to be 282730 joules per mole at 20° C. From an assessment of the maximum error which might occur in each quantity measured, it is estimated that the value 282730 should be right to within 3 parts in 1000.

The result for the heat of reaction when nitrous oxide burns in carbon monoxide is 364340 joules per mole. A discussion of the limiting accuracy in each factor leads to a suggestion that this result also is reliable to 2 or 3 parts in 1000.

From these two results, the heat of formation of N_2O at constant atmospheric pressure and at a temperature of 20° C. is found to be 81610 joules, *i.e.*, $19 \cdot 5_0$ kg.-cal.$_{15}$ per mole. This quantity is determined as the difference between two heats of combustion, which were carried out with the same apparatus used in the same manner for both sets of experiments.

A Bomb Calorimeter Determination of the Heats of Formation of Nitrous Oxide and Carbon Dioxide.

By R. W. Fenning, M.B.E., B.Sc., D.I.C., and F. T. Cotton, B.Sc.,
Engineering Department, National Physical Laboratory, Teddington.

(Communicated by Sir Joseph Petavel, F.R.S.—Received February 3, 1933.)

In the course of a specific heat investigation at the National Physical Laboratory the need arose for an accurate knowledge of the heat of formation of nitrous oxide and since the published values of this constant differed widely it was decided to make a new determination.

The work was undertaken by both the Physics and Engineering Departments of the National Physical Laboratory, and, whereas the former used a continuous flame method* involving combustion at constant pressure, the latter adopted the explosion method and a constant volume reaction. It is the purpose of this paper to give an account of the results obtained by the explosion method.

The mode of attack is already familiar and consisted, in this investigation, in exploding mixtures of the approximate composition $3CO + 2N_2O$ and $3CO + O_2$ and ascribing the difference in their heats of reaction to the splitting up of the $2N_2O$ into $2N_2$ and O_2. Subsequently a series of experiments was made on the mixture $3H_2 + 2N_2O$ and its heat of reaction was compared with that of the calibration mixture $3H_2 + O_2$. The result confirmed the value already obtained from the CO experiments.

A reference has just been made to the calibration mixture and it should be pointed out at once that the heat capacity of the calorimeter was determined by exploding $3H_2 + O_2$ mixtures in it, and hence the values given for the heats of formation of N_2O and CO_2 are based on the heat of formation of H_2O (liquid) which has been taken as 68,320 gm. $\text{cal.}_{15°}$ at constant pressure and at 25° C.

The accuracy with which the required heat of formation can be determined depends on the precision with which measurements can be made of the masses of the gases taking part in the reaction and the quantities of heat evolved. For the measurement of the latter the bomb calorimeter shown in fig. 1 was specially designed by one of us (R. W. F.). This calorimeter is of the "adiabatic type" and, whilst its design was inspired by the valuable work and experience of previous investigators, new features have been introduced to meet the particular experimental conditions under which a comparatively small quantity of heat is generated in a very short interval of time. The quantity of heat being small it was necessary to keep down the heat capacity of the calorimeter

---

* Awbery and Griffiths, p. 1.

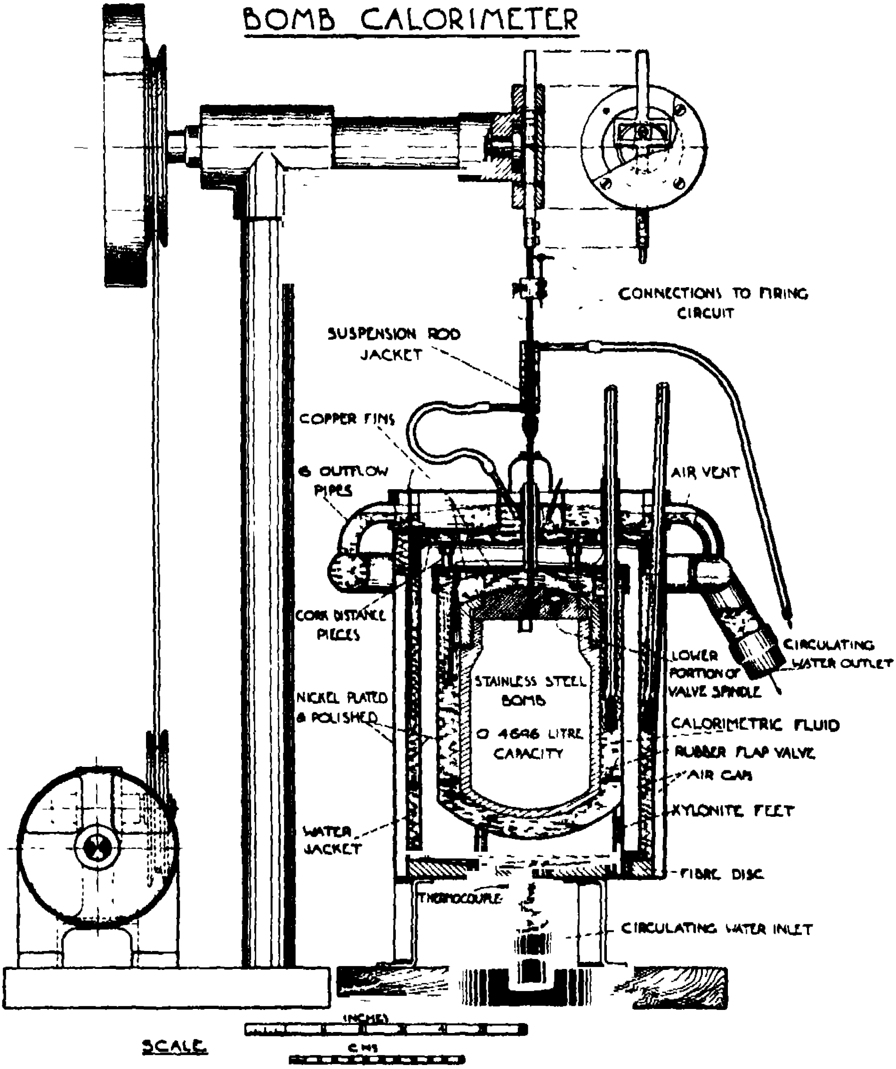


FIG. 1.

and this was achieved by dispensing with the usual propeller for stirring the calorimetric fluid and by devising a valve most of which could be removed from the bomb without releasing the bomb contents.

For stirring the calorimetric fluid the bomb itself was used. A light brass fitting, which divided the water space vertically into two parts and also carried horizontal flap valves opening upwards in one half and downwards in the other, was attached to the bomb. The bomb was given a regular and continuous reciprocating motion and this set up a circulation of the calorimetric fluid which appeared to result in quick and thorough mixing. As will be seen from fig. 1, a piece of thin walled brass tube was used to connect the bomb to the external reciprocating gear, and this tube and an insulated wire within it formed part of the circuit for igniting the charge.*

The rapid rise in temperature of the calorimetric fluid, to which allusion has already been made, called for equally rapid means of varying the temperature of the thermal jacket so that the temperature balance could be maintained and the adiabaticity of the calorimeter preserved. Instead of attempting to heat the water in the jacket, however, it was decided to replace it by water at the required temperature. By circulating water through the jacket at the rate of about 4 gallons per minute, the whole of the relevant water can be renewed in about 3 seconds, and hence, given two sources of water at different temperatures and a " mixing valve,"† it is possible to vary the temperature of the jacket from that of one source to that of the other in a few seconds. The installation consisted of two large tanks, each provided with a rotary pump for circulating the water through a pipe line passing *close to the entry* to the calorimeter jacket. The water in one tank was kept at the temperature of the calorimetric fluid before the ignition of the bomb charge, whereas the temperature of the water in the other tank corresponded to the temperature of the calorimetric fluid after the ignition of the charge. For the half-hour or so prior to ignition, the water was taken entirely from the cooler tank and returned thereto. During this period the temperature of the calorimetric fluid rose slowly—about 0·002° C. per minute—owing to the work done in stirring, etc., and the temperature of the water in the tank was adjusted accordingly. The two mercury-in-glass thermometers shown in fig. 1 were used to effect the balance, which was generally maintained within a few thousandths of a degree centigrade.

For the quick-rise period following ignition the temperature balance was sought by endeavouring to maintain at zero the galvanometer reading given by three differential thermocouples in series, these thermocouples being distributed in the jacket water and calorimetric fluid respectively. The maximum divergence from the zero averaged about 0·14° C. and since any divergence in one direction was roughly neutralized at the time by imposing a divergence in the opposite direction, the heat exchange between calorimeter and jacket during this short period of about 1 minute was probably negligible.*

---

* The charge was fired by fusing a short length of 0·047 mm. diameter platinum wire.

† The " mixing valve " consisted of two screw-down valves having their spindles geared together.

At the end of this period the rate of temperature rise was low enough to permit the balance being closely maintained by means of the mercury thermometers.

This brief† account of special features of the calorimeter would be incomplete without reference to the water jacket (syphon action) on the bomb suspension rod for preventing heat transfer along this rod, and to the copper fins on the covers whose chief function is to promote the rapid equalization of the temperature of the cover with that of the fluid directly below it.

In carrying out an experiment, the calorimeter and jacket thermometers were read alternately at two-minute intervals during the periods prior, and subsequent, to ignition when the heat generated was due to the stirring of the calorimetric fluid. Between these two periods, the heat of the reaction was communicated to the calorimeter and readings were then taken at shorter intervals. A plot of the calorimeter thermometer readings is given in fig. 2, and it is noticeable how closely these readings conform to a straight line during the stirring periods. The thermometers, which were made by B. Black & Son, Ltd., of London, were graduated in hundredths of a degree centigrade and were subjected to a routine method of tapping before each reading was taken.

Throughout the experiments recorded in this paper the mass of the reacting gases was selected so as to give sensibly the same temperature rise—about 2·73° C.—of the calorimetric fluid and since the charges were fired at the same initial temperature (20·16° C.) the thermometer was always used over practically the same interval on its scale. This temperature interval was the one used when determining the "water equivalent" of the calorimeter by explosions of a $3H_2 + O_2$ mixture and hence the absolute value of the interval is of no consequence.

Having dealt with the method of measuring the reaction heats, a brief reference will now be made to the means of determining the mass of the reacting gases. Since, in the mixtures, the carbon monoxide or the hydrogen was always in considerable excess of the quantity that could combine with the oxygen or the nitrous oxide, careful measurements of the latter gases only

---

* It was found by experiment that a difference of 1° C. between calorimetric fluid and the outer jacket, caused the temperature of the former to vary at the rate of 0·005° per minute. Hence if the uncompensated divergence was less than 0·05° C. for half a minute, the resulting heat transfer would correspond to a temperature variation of the calorimetric fluid of less than 0·000125° C.

† For a more detailed account see 'Aero. Res. Ctee. Rep. and Mem.,' No. 1513 (*in course of publication*).

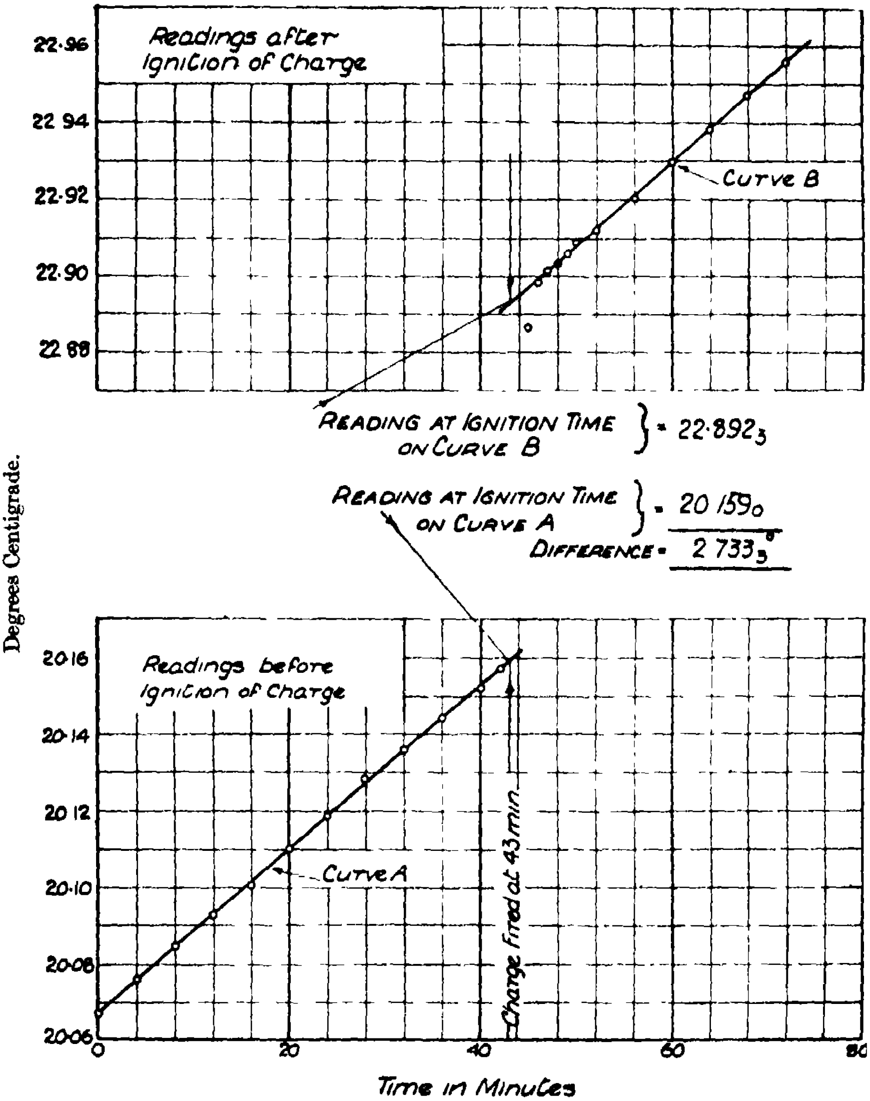


Fig. 2.—A temperature-time graph of the calorimetric fluid. Mixture $3H_2 + 2N_2O$.

were required. The bomb was placed in a water bath at about room temperature and charged with oxygen to about 980 mm. (absolute) or nitrous oxide to about 1500 mm. (absolute), the pressure readings being taken on a mercury

U-tube. The volume of the bomb was known and hence pressure, volume and temperature readings were available from which to determine the mass of gas. For oxygen the density at 0° C. and 760 mm. was taken as 1·42904 gm. per litre and PV/T was assumed to be constant over the small range of pressure concerned. Wth nitrous oxide, however, this constancy could not be assumed for the pressure range was considerably greater, and this gas was known to depart considerably from the perfect gas laws. From a survey of the work of Leduc and Sacerdote, Berthelot, Rayleigh and Batuecas, the following representative values of the mean coefficient of compressibility* per centimetre (Hg) were deduced for the pressure range two to one atmospheres . $81 \cdot 0 \times 10^{-6}$, $77 \cdot 4 \times 10^{-6}$, $76 \cdot 2 \times 10^{-6}$ and $74 \cdot 6 \times 10^{-6}$ at 16·35° C., 20·1° C., 21·4° C., and 23·1° C. respectively. These were used to determine the volume of the nitrous oxide when its pressure was reduced isothermally to one atmosphere, and a coefficient of expansion of 0·00375 was then employed for the determination of the volume at 0° C. The mass of the normal litre of N_2O was taken as 1·9778 gm.†

The evaluation of the mass of oxygen or nitrous oxide involves also a knowledge of the purity of these gases, and of the O_2 content of the CO or H_2. The purity of the oxygen, ignoring its water vapour content, was found to be 99·75 per cent.‡ As the gas was taken from a high pressure container and precautions were taken to keep it dry, the water vapour content was probably less than 1 part in 4000 and no allowance was made for it.§

The nitrous oxide and its analysis were kindly supplied by Messrs. Coxeter & Son, Ltd., London. The analysis from the gas phase gave :—

Nitrous oxide, more than	99·98 per cent.
Nitrogen and oxygen, less than	0·01 ,,
Carbon monoxide	Absent.
Carbon dioxide, less than	0·005 per cent.
Water vapour, less than	0·005 ,,

---

* The mean coefficient of compressibility between p' and $p = \dfrac{1 - \dfrac{p'v'}{pv}}{p' - p}$.

† Int. Crit. Tables.

‡ The 0 25 of impurities included about 0·035 of H_2.

§ If the oxygen is less pure than the figure given, then the water equivalent of the calorimeter will have been overestimated and the heat of combustion of N_2O in CO or H_2 will be too high. It should be noted, however, that the N_2O has been taken as pure, whereas from the analysis the impurities might amount to almost 1 part in 5000. If the impurities were of this order their neglect would practically neutralize the possible 1 in 4000 water-vapour effect.

The firm mentioned use an interferometer for their analytical work and their results are accepted as having a high degree of accuracy. A check test for oxygen was, in fact, made but no trace of that gas was detected.

The carbon monoxide used was part of a consignment obtained some two years ago in connection with another investigation. It was kindly supplied by the Chemical Research Laboratory. Its composition was :—

Carbon monoxide	99·55	per cent.
Carbon dioxide	0·20	,,
Nitrogen	0·25	,,

The hydrogen was tested for oxygen and a trace of that gas was found to the extent of about 0·005 per cent. The total impurities in the hydrogen were less than 0·1 per cent.

During the combustion of the gas mixtures it is assumed that the O_2 burns to CO_2 or H_2O and the N_2O splits up into nitrogen and oxygen and the latter combines with the CO or H_2. This is not a pure assumption since the products of representative mixtures were tested for nitric oxide, but no appreciable trace of that gas appeared to be present.

The experiments for determining the heat of formation of nitrous oxide can be divided into three main series. The first series gave a value of − 19,938 gm. $\text{cal.}_{15°}$ per mol. at constant volume and at 20° C. The results were quite consistent amongst themselves, but the series was not considered to be quite satisfactory since the temperature response of the water jacket was not sufficiently rapid. To hasten the response, the mixing valve was brought closer to the inlet to the calorimeter jacket, the water capacity of the latter was reduced* and the rate of water circulation was increased. Further, the single differential thermocouple that had been used previously was replaced by three such thermocouples in series, the junctions being distributed in the jacket water and calorimetric fluid respectively. These modifications, which are included in fig. 1, gave the desired result and a second series of experiments was put in hand. This series comprised . (i) three explosions of the approximate mixture $3CO + O_2$, (ii) four explosions of the approximate mixture $3CO + 2N_2O$, of which one had eventually to be discarded owing to an unfortunate omission in booking one of the charging readings ; (iii) three explosions of the approximate mixture $3H_2 + O_2$.

---

* This was effected by inserting a fibre disc in the bottom of the jacket and forming an air gap round the cylindrical portion, see fig. 1. The fibre disc is essentially an improvization, and in any re-design it would be superseded by a suitable shaping of the base of the outer vessel of the calorimeter and some equivalent modification in metal.

As the three $3H_2 + O_2$ experiments were carried out for the purpose of obtaining the "water equivalent" of the calorimeter, these results will be given first. They were :—

		Deviation from mean.
Water equivalent in gm.$_{15°}$ =	1241·29	−0·31
	1241·58	−0·02
	1241·94	+0·34

The mean value is 1241·60 gm.$_{15°}$ and the maximum deviation from the mean is 1 in about 3700. The "95·45 per cent. probable error" of the mean is ± 0·37 or ± 1 in about 3350. In computing the water equivalent the heat of formation of H_2O (liquid) at constant pressure and at 25° C. was taken as 68,320 gm. cal.$_{15°}$, *i.e.*, 67,486 gm. cal.$_{15°}$ at constant volume and at 20° C. Roth, in his critical survey,[*] gives 68,330 and quite recently Rossini[†] obtained 68,313[‡] ± 10, both being constant pressure values at 25° C.

From the three $3CO + O_2$ explosions, the undermentioned values of the heat of formation of CO_2 from CO and O_2 were obtained :—

		Deviation from mean.
Heat of formation of CO_2 at constant volume and at 20° C. in gm. cal.$_{15°}$ =	67,393	+29
	67,351	−13
	67,349	−15

The mean is 67,364 per mol. of CO_2 and the maximum deviation from the mean is 1 in 2300. The "95·45 per cent. probable error" of the mean is ± 29. Combining this probable error with that in the water equivalent, given above, by taking the root of the sum of their squares, the probable error in the 67,364 figure becomes ± 35.

The values given by the $3CO + 2N_2O$ explosions for the heat of combustion of N_2O in CO were :—

		Deviation from mean.
Heat of combustion at constant volume and at 20° C. in gm. cal.$_{15°}$ =	87,447	+46
	87,399	− 2
	87,358	−43

---

[*] 'Z. Elektrochem.,' vol. 26, p. 288 (1920).

[†] 'Bur. Stand. J. Res.,' vol. 6, p. 37 (1931).

[‡] This would be 68,296 on the basis of the conversion factors used in this paper, see p. 26.

The mean is 87,401 per mol. of N_2O and the maximum deviation from the mean is 1 in about 1900. The " 95·45 per cent. probable error " is ± 51 and the combination of this error with that in the water equivalent gives ± 58.

The difference between the values obtained from the last two sets of experiments gives the heat of formation of nitrous oxide. Hence

The heat of formation of nitrous oxide at constant volume and at 20° C. .... $-$ 20,037 gm. cal.$_{15^\circ}$.

The probable limits of error, on the basis of those already given, are about ± 60, but a review of the precision with which the various measurements can be made suggests that it would be advisable to increase this figure to ± 70.

The third series of experiments comprised: (i) a single explosion of a $3H_2 + O_2$ mixture to check the constancy of the " water equivalent " reading and hence the constancy of the temperature interval on the thermometer scale; and (ii) three explosions of a $3H_2 + 2N_2O$ mixture.

According to the single $3H_2 + O_2$ explosion :—

The water equivalent in gm.$_{15^\circ}$ = 1241·53.

Since this confirmed the previous value of 1241·60 within 1 part in 17,000, there appeared to be no grounds for departing from the former value.

The three $3H_2 + 2N_2O$ explosions gave the following results :—

		Deviation from mean.
Heat of combustion per mol. of N_2O at constant volume and at 20° C. in gm. cal.$_{15^\circ}$ =	87,521	+1
	87,519	−1
	87,520	0

The mean value is 87,520 and the maximum deviation from the mean happens to have the extremely small value of 1 in about 70,000. The " 95·45 per cent. probable error " of the mean is ± 1·3 and this becomes ± 26 when combined with the " water equivalent " error (p. 24).

If the result of this set of experiments is taken in conjunction with the $3H_2 + O_2$ calibration experiments in which 67,486 gm. cal.$^{15^\circ}$ was taken as the heat of formation of H_2O (liquid), then

$$3H_2 + 2N_2O = 2H_2O + H_2 + 2N_2 + 2 \times 87{,}520 \text{ gm. cal.}_{15^\circ}$$

and

$$3H_2 + O_2 = 2H_2O + H_2 + 2 \times 67{,}486 \text{ gm. cal.}_{15^\circ}$$

and therefore

The heat of formation of 1 mol. of N_2O at constant volume and at 20° C. .... = – 20,034 gm. $cal._{15°}$.

This value is in such close agreement with the – 20,037 given by the CO experiments, that the fears of possible incomplete combustion of the 3CO + $2N_2O$ mixtures, owing to their dryness and the absence of hydrogen, appear to have been groundless. Although the probable limits of error in the –20,034 quantity appear to be only ± 26, it is considered advisable as before to widen them to ± 70.

Collecting the results it may be concluded from the above experiments that:—

(i) The heat of formation of nitrous oxide at constant volume and at 20° C. is – 20·03 k. $cal._{15°}$, which is probably correct to within ± 0·07 k. $cal._{15°}$.

(ii) The heat of formation of CO_2 (from CO and O_2) at constant volume and at 20° C. is $67·36_4$ ± 0·035 k. $cal._{15°}$.

(iii) The ratio of the heat of formation, at constant volume and at 20° C., of H_2O (liquid) to that of CO_2 is 1·0018.

The conversion of these constant volume determinations to constant pressure values gives :—

(iv) The heat of formation of N_2O at a constant pressure of 1 atmosphere and at 20° C. as – 19·74 ± 0·07 k. $cal._{15°}$ or – 82,600 ± 290 international joules.*

(v) The heat of formation of CO_2, from CO and O_2, at a constant pressure of 1 atmosphere, and at 20° C. as $67·65_5$ ± $0·03_5$ k. $cal._{15°}$ or 283,090 ± 150 international joules.

The tabular statements given below enable these values to be compared with those obtained by other investigators.

---

* Taking 1 gm. $cal_{15°}$ = 4·186 absolute joules and 1 international joule = 1·00039 absolute joules.

For nitrous oxide :—

Investigator.	Date.	Method.	Heat of formation of N_2O at constant pressure.
Berthelot	1880	Combustion at constant volume of CO and N_2O	− 20·6 k cal. per mol.
Thomsen	1882	Combustion at constant pressure of N_2O in H_2 Combustion at constant pressure of CO in N_2O	− 18·01 k cal.$_{19^\circ}$ per 0·999 mol. (at 19° C.). − 17·47 k. cal $_{17^\circ}$ per mol. (at 17° C.)
Sutton .	1932	Combustion at constant volume of N_2O in H_2	20·5 ± 0·3 k cal $_{15^\circ}$ per mol. (at 20° C.)
Awbery and Griffiths	1933	Combustion at constant pressure of N_2O in CO	− 19·5 k cal $_{15^\circ}$ per mol. (at 20° C.)
Fenning and Cotton	1933	Combustion at constant volume of N_2O in CO Combustion at constant volume of N_2O in H_2	− 19·74$_6$* − 19·74$_3$* Value taken − 19·74 ± 0·07 k cal $_{15^\circ}$ per mol. (at 20° C.).

For the heat of formation of CO_2 from CO and O_2 :—

Investigator.	Date.	Method.	Heat of formation at constant pressure per mol.
Berthelot	1878	Combustion at constant pressure.	68·17 k cal.
Berthelot	1880	Combustion at constant volume	68·28 k cal.
Thomsen	1882	Combustion at constant pressure.	67·96 k. cal.$_{19^\circ}$ (at 19° C.).
Berthelot and Matignon	1893	Combustion at constant volume.	68·2 k cal.
Rossini	1931	Combustion at constant pressure.	67·615† ± 0·030 k. cal.$_{15^\circ}$ (at 20° C.).
Roth and Banse	1932	Combustion at constant volume.	67·8$_7$ ± 0·1 k. cal $_{15^\circ}$ (at 20° C.).
Awbery and Griffiths	1933	Combustion at constant pressure	67·57 k. cal.$_{15^\circ}$ (at 20° C.).
Fenning and Cotton	1933	Combustion at constant volume.	67·65$_8$* ± 0·03$_8$ k cal $_{15^\circ}$ (at 20° C.).

* Based on 68·320 k. cal.$_{15^\circ}$ as the heat of formation of H_2O (liquid) at constant pressure and at 25° C.

† *This would be 67·598 on the basis of the conversion factors used in this paper (see footnote * on previous page).*

In conclusion it may be mentioned that the investigation has been free from any combustion difficulties and the bomb calorimeter has given very consistent results. In fact the accuracy of the N_2O determination appears to depend mainly on the correctness of the gas analysis and on the accuracy of the figures taken for the compressibility. An error in the oxygen analysis affects the N_2O result but not the CO_2 result, for the same oxygen was used in the $3CO + O_2$ experiments as in the $3H_2 + O_2$ calibration experiments.

Acknowledgments.

This investigation was financed by the Department of Scientific and Industrial Research and carried out under the scientific supervision of the Executive Committee of the National Physical Laboratory.

The authors desire to thank the Chemical Research Laboratory for supplying the carbon monoxide and Messrs. Coxeter & Son, Ltd., for preparing and analysing so pure a sample of nitrous oxide.

Summary.

A re-determination was required of the heat of formation of nitrous oxide. By means of a bomb calorimeter, designed for the purpose, the heats of reaction of oxygen and nitrous oxide in carbon monoxide and hydrogen respectively were measured. Thus, by difference, two values were obtained for the heat of formation of nitrous oxide. Incidentally a value for the heat of formation of CO_2 from CO and O_2 was also obtained. Since the combustion of oxygen in hydrogen formed the calibration process, all the values given are based on the acceptance of 68,320 gm cal.$_{15^\circ}$ as the heat of formation of H_2O (liquid) at constant pressure and at 25° C.

On this basis, the heats of formation of N_2O and CO_2 at a constant pressure of 1 atmosphere and at 20° C. were found to be $-19 \cdot 74 \pm 0 \cdot 07$ k. cal.$_{15^\circ}$ or $-82{,}600 \pm 290$ international joules, and $67 \cdot 65_5 \pm 0 \cdot 03_5$ k. cal.$_{15^\circ}$ or $283{,}090 \pm 150$ international joules respectively.

The Upper Pressure Limit in the Chain Reaction between Hydrogen and Oxygen.

By G. H. GRANT and C. N. HINSHELWOOD, F.R S.

(Received March 4, 1933.)

Mixtures of hydrogen and oxygen above 450° C. possess two critical pressure limits between which the normally slow combination of the gases becomes explosive. The most satisfactory theory attributes the explosion to the "branching" of reaction chains, and explains the existence of the lower limit by the deactivation of chain carriers at the wall of the vessel, and the upper limit by deactivation in the gas phase.* At the limits one or other of these deactivation processes is just rapid enough to balance the multiplication of the chains. The conditions governing the position of the lower limit have been investigated in some detail and found to be accounted for fairly well by the simple theory.†

With regard to the upper limit, while several investigations have left little room for doubt that it is governed by gas phase deactivation,‡ the exact nature of this process has not hitherto been satisfactorily worked out. A more complete investigation of the problem therefore seemed to be called for. In order to make this as thorough as possible, certain alternative theories§ have first been reconsidered, namely those which make the limit a function of surface adsorption relationships, and those which make it depend on various thermal conductivities. But the results show that neither the surface of the vessel, nor the thermal conductivity of the gas are fundamentally important in determining the limit, which proves to depend upon ternary collisions in the gas phase. On the basis of the latter hypothesis an approximate quantitative theory can be constructed.

---

* For summary and references to previous work and to the general chain theory of Semenov see Garstang and Hinshelwood, 'Proc. Roy. Soc.,' A, vol. 130, p. 640 (1931).

† Hinshelwood and Moelwyn-Hughes, 'Proc. Roy. Soc.' A, vol. 138, p. 311 (1932).

‡ Garstang and Hinshelwood, *ibid.*, vol 134, p. 1 (1931).

§ For summary see 'Trans. Faraday Soc.,' vol. 28, p. 184 (1932).

Table I.

Temperature 550° C
Alumina bulb.

H_2/O_2.	P_e.
0·28	149
0·50	127
0·67	111
1·0	101
1·5	93
2·0	83
3·0	71

Temperature 550°
Second alumina bulb

H_2/O_2.	P_e
0·26	153
0·50	128
0·71	151
1·03	104
1·51	95
1·99	84

Temperature 550°
Silica bulb.

H_2/O_2.	P_e
0·25	165
0·50	143
0·67	128
1·0	116
1·5	104
2·0	97
3·0	91

Temperature 552° C.
Second silica bulb

H_2/O_2.	P_e
0·25	168
0·50	141
1·0	125
1·5	119
2·0	111
2·5	108
3·0	107

Table II.

Ratio H_2/O_2 = 1·1.
Alumina bulb.

Temperature	P_e.
579	187
565	135
550	101
535	77
520	58·5
500	37·5

Ratio H_2/O_2 = 1·1
Second alumina bulb

Temperature	P_e
580	184
565	141
550	104
535	79
520	57·5
500	39

Ratio H_2/O_2 = 1·1
Silica bulb

Temperature	P_e
565	158·5
550	120
535	88
518	64
500	44

Ratio H_2/O_2 = 1·1.
Second silica bulb.

Temperature	P_e.
568	154
552	125
537	100
522	77
507	58

Influence of the Vessel Surface.

The upper limit is determined by mixing suitable proportions of the gases at relatively high partial pressures in the reaction vessel, and then gradually evacuating.* As the limit is passed an explosion takes place accompanied by a kick on the manometer, and sometimes by a click and a flash. For the method to give correct results the rate of the heterogeneous reaction, which is always occurring to some extent, must not be great enough to produce appreciable quantities of water during the time of the experiment. Water lowers the explosion pressure and may even inhibit the explosion entirely.† Under ordinary conditions the results in a silica vessel are not affected by this factor.

Tables I, II and III give the upper explosion pressure, P_e in different vessels of alumina and silica for various temperatures and proportions of gases. The relation of P_e to temperature and to the ratio H_2/O_2 is exactly the same in the two kinds of bulb as shown in figs. 2 and 3. At a given temperature there is a small difference between the absolute P_e values for alumina and silica vessels. This, however, can be understood. Although most of the deactivation occurs in the gas phase, a certain contribution from surface deactivation—which is all important at the lower limit—must still exist and will account for minor variations. A still more important factor is the greater surface activity of alumina in producing water. The influence of this appeared most clearly in some experiments made with a vessel of crude zirconia. At first for a 1 : 1 mixture at 550°, P_e was 48 mm., but the speed of the surface reaction was observed to be great enough to give appreciable amounts of water during working. By progressively increasing the speed of manipulation values up to 92 were found compared with 104 for alumina. This factor cannot operate in quite the same way with the alumina, since the results given in the tables were in fact uninfluenced by considerable variations in the speed of working. At the same time, the condition of the surface may easily control the moisture content of the gas to a small extent, and this will cause some degree of variation in the explosion pressure. In one series of experiments an

Table III.

Diameter of silica bulb	7·5 cm	4·9 cm	1·8 cm
P_e for 1 : 1 mixture at 552°	125 mm	125 mm	126 mm.

---

* Thompson and Hinshelwood, 'Proc. Roy. Soc.,' A, vol. 122, p. 610 (1929).

† Garstang and Hinshelwood, *loc. cit.*

alumina bulb began by giving results absolutely identical with those for silica, but soon relapsed to the slightly lower ones recorded in the table.*

There can be little doubt that the differences between the various surfaces are due to secondary causes. The general similarity is much more significant than the small variations. This is evidence against theories which make the upper limit represent a point where discontinuities in adsorption first allow suitable chain carriers to escape from the wall. The experiment† in which streams of hydrogen and oxygen will only ignite when a solid rod is interposed indicated, it is true, that something necessary for the chains may in the first instance have to be formed at the wall. But the condition for explosion is that the rate of branching of the original chains shall exceed the rate of deactivation. This condition has nothing to do with the initial formation of the chain carriers.

Influence of the Thermal Conductivities.

According to one view, the explosion might be caused by the intense local heating of certain very active centres on the surface. Our attention was again drawn to this possibility by the absence of the explosion phenomenon in a silver vessel.‡ Though this observation can quite well be explained on other grounds it might have meant simply that owing to the high thermal conductivity of the metal no local departures from isothermal conditions occurred. Apart from the fact that the local heating theory cannot be used to explain the effect of vessel diameter on the lower limit, it now appears unhelpful for the following reasons It would lead to the prediction of far greater differences between silica, alumina and zirconia vessels than are in fact found, and it also fails to account for the variation of P_e with the composition of the gas mixture. The upper limit would be the point at which the cooling power of the gas was great enough to keep the temperature of the active points from rising. Experiments were made from which curves could be plotted showing the amounts of electrical energy required to keep a given tungsten filament at a known temperature (below the reaction temperature) in mixtures of hydrogen and oxygen at any pressure and of any composition. From these curves it was found that the relative pressures required to produce equal cooling bore no resemblance to the relative values of P_e, as the following numbers show :—

---

* In our opinion variations in the explosion pressure with "pre-treatment" of the surface are probably due essentially to changes in the amount of water introduced into the gases. In one of the alumina bulbs P_e was found to fall until the gas refused to explode. The rate of the heterogeneous reaction was now found to have become considerable. Up to this point hydrogen had been added to the vessel before oxygen. On reversing the procedure the surface reaction was found to be slow, and normal concordant results for P_e were once more obtained. With silica or an inactive alumina no variations in P_e with order of addition were found, beyond those which could be accounted for by changes in the ratio $H_2 : O_2$ owing to the "dead-space" of those vessels with rather wide necks.

† Alyea and Haber, 'Z. phys. Chem.,' B, vol. 10, p. 193 (1930).

‡ Hinshelwood, Moelwyn-Hughes and Rolfe, 'Proc. Roy. Soc.,' A, vol. 139, p. 521 (1933).

Mixture	$2H_2 : O_2$	$H_2 : O_2$	H_2 $2O_2$
Relative pressure of gas for equal cooling	100	350	980
Relative value of P_e	100	119	147 (silica)
	100	123	149 (alumina)

Even making full allowance for variations in accommodation coefficients of different kinds of surface, it seems clear that no simple theory of the variation of P_e with composition can be developed on these lines. The theory of ternary collisions, however, provides an adequate explanation. This will now be considered.

Influence of Inert Gases and of Hydrogen-Oxygen Ratio on the Explosion Pressure : Theory of the Gas Phase Deactivation.

Inert gases lower the partial pressure of hydrogen and oxygen at the limit, and hence contribute to the deactivation. Moreover, helium, the molecular velocity of which is high, is much more effective than argon. Furthermore, with excess of hydrogen the limit is lower than with excess of oxygen. These facts suggest that the deactivating process may be of the nature of a ternary collision, and that the relative speeds of the different types of molecule largely determine their effectiveness. The following simple theory accounts more or less quantitatively for many important facts about the upper limit.

Let X and Y be two kinds of molecule or atom which collide in the course of a chain, and let there be a probability ν that at such a collision the chain branches. If, however, any third molecule arrives while X and Y are associated together in the " collision complex," then the branching is prevented. At the pressure where explosion just ceases to be possible the rate of branching is balanced by the rate of deactivation in the ternary collisions. Thus

$$\nu k[X][Y] = Z_{H_2}[X][Y][H_2] + Z_{O_2}[X][Y][O_2] + Z_M[X][Y][M],$$

where M represents an inert gas present.

k is the number of encounters of X and Y at unit concentration of each, and the Z factors are proportional to the collision frequencies of the temporary complexes XY with molecules of hydrogen, oxygen, and the inert gas respectively. At the limit, therefore, we have

$$Z_{H_2}[H_2] + Z_{O_2}[O_2] + Z_M[M] = \nu k. \tag{1}$$

From equation (1) it follows that for a given ratio of hydrogen to oxygen, the sum of their partial pressures should decrease linearly with the amount of inert gas added. Table IV and fig. 1 show that this relation holds.

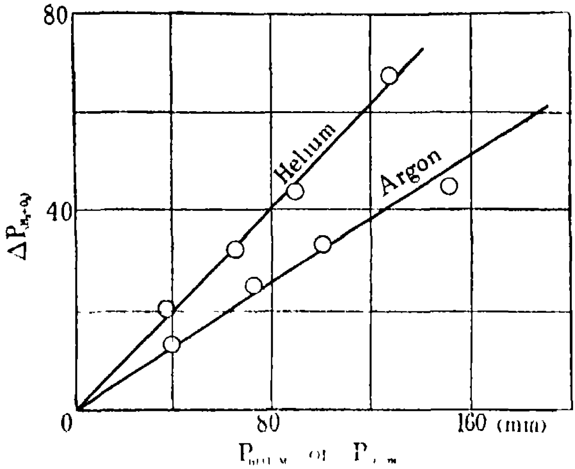


Fig. 1.—Variation of partial pressure with inert gas, Table IV.

Table IV.—Influence of Helium and Argon.

$H_2 : O_2 = 1 : 1$, temperature 554° C.

$P_{H_2+O_2}$.	P_{Helium}.	$\Delta P_{H_2+O_2}$	$P_{H_2+O_2}$.	P_{Argon}	$\Delta P_{H_2+O_2}$
133	0	0	133	0	0
113	37	20	119	39·5	13
101	65	32	108	71·5	25
89	89	44	99·5	100	33·5
66	127	67	88	151	45

Each of the above figures is the mean of three determinations. The ratio of the slopes for $\Delta P_{H_2+O_2}/[He]$ and $\Delta P_{H_2+O_2}/[A]$ is 1·6.

A corresponding series of experiments with a ratio $H_2 : O_2 = 1 : 2$ gave $\Delta P/[He] : \Delta P/[A] = 1{\cdot}54 : 1$. ([He] is the same as P_{Helium}.)

The slopes should be proportional to Z_{He} and Z_A respectively, a point which will be returned to later.

Equation (1) also shows that in the absence of an inert gas

$$P_{H_2} = \text{Constant} - b \, . \, P_{O_2},$$

where $b = Z_{O_2}/Z_{H_2}$.

This linear relation also holds over a considerable range as shown in fig. 2 and in Table V.

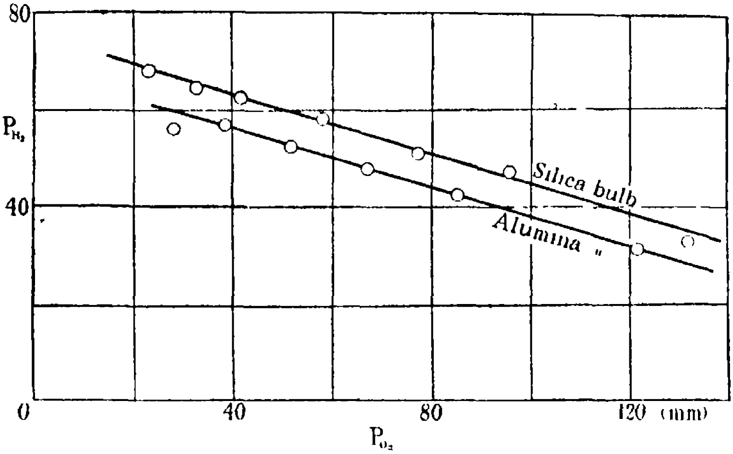


FIG. 2.—Variation of P_{H_2} with P_{O_2}, Table V.

Table V.

Ratio $H_2 : O_2$.	P_s (calc.).	P_s (obs.).
1 : 4	165	165
1 : 2	138	143
2 : 3	127	128
1 : 1	115	116
3 : 2	104	104
2 · 1	98	97
3 1	91	91

The calculated values are those given by the equation $P_{H_2} = 76 - 0 \cdot 325 \, P_{O_2}$, P_s being $P_{H_2} + P_{O_2}$. For the second alumina bulb the equation expressing the results ran $P_{H_2} = 68 - 0 \cdot 30 \, P_{O_2}$. The best value of b from all the results is not far from that of Table V, the average being 0·33.

The values of Z, the ternary collision number for unit concentrations, depend upon the molecular diameters and speeds. For a ternary collision between a molecule M and the pair XY it can easily be shown* that Z_M is proportional to $\sigma^2_{XYM}/\mu^{\frac{1}{2}}_{XYM}$, where σ_{XYM} is the sum of the molecular diameters of the complex XY and the colliding molecule M. $\mu_{XYM} = \frac{m_{XY} m_M}{m_{XY} + m_M}$, m_M, etc., being the molecular weights. From the form of μ, it appears that the mass of the third molecule only plays an important part if the mass of the pair XY is considerably greater than that of hydrogen or helium. Since Z_{H_2} is obviously much greater than Z_{O_2} and Z_{He} is greater than Z_A, we infer that the mass of XY is relatively large. Given this, the calculation is not very sensitive to the exact value assumed for the mass. We shall not be far wrong in assuming therefore that it is approximately the same as the mass of $H_2 + O_2$. We also take $\sigma_{XY} = \sigma_{H_2} + \sigma_{O_2}$. Using the values for $\sigma \times 10^8$ given in Tolman's "Statistical Mechanics," namely $\sigma_A = 2{\cdot}84$, $\sigma_{He} = 1{\cdot}89$, $\sigma_{O_2} = 2{\cdot}93$ and $\sigma_{H_2} = 2{\cdot}36$ we find that $Z_{O_2}/Z_{H_2} = 0{\cdot}39$ compared with the value $0{\cdot}33$ of the constant b in the above formula.

Furthermore, Z_{He} should be $1{\cdot}77$ times as great as Z_A, while actually helium is found to be about $1{\cdot}6$ times as effective as argon.

The actual slope of the curve for $\Delta P_{H_2+O_2}$ against inert gas concentration should be $2Z_M/(Z_{H_2} + Z_{O_2})$ for the 1 : 1 hydrogen-oxygen mixture. The observed value for argon is appreciably smaller, namely $0{\cdot}32$ compared with the calculated $0{\cdot}52$. But the cross comparison between the monatomic and the diatomic gases cannot be expected to be as satisfactory as that of the monatomic gases or the diatomic gases among themselves.

Having regard to uncertainties in the molecular diameters the general agreement must be regarded as rather close, and shows that the relative speeds of the different molecules is one of the principal factors† determining the variation of the upper limit with composition. Thus the theory of the gas phase deactivation of a fairly massive binary complex in a ternary collision receives strong support.

Influence of Temperature.

Since variation of Z with temperature is small, equation (1) shows that the temperature coefficient of the upper explosion pressure will be practically that of ν. The probability of branching presumably depends upon the energy of the colliding pair; thus P_e should vary with temperature according to an Arrhenius equation. The degree to which this is satisfied is shown in fig. 3. The "energies of activation" derived from the two sets of measurements plotted in the curves are respectively 25,500 calories (silica bulb) and 25,800 calories (alumina bulb). These values were both found with a ratio $H_2 : O_2 = 1 : 1$. A series of experiments made some years ago in a silica bulb with a ratio $H_2 : O_2 = 2 : 1$ gave on plotting in a similar way a value 26,500 calories.

---

* For methods of calculation see, for example, Tolman, "Statistical Mechanics," and especially Steiner, 'Z. phys. Chem.,' B, vol. 15, p. 249 (1932).

† This does not apply to certain substances such as water and the halogens which have a powerful *specific* action.

If the branching probability varies exponentially with temperature, as now appears, it remains to be explained why the lower limit is so nearly constant.

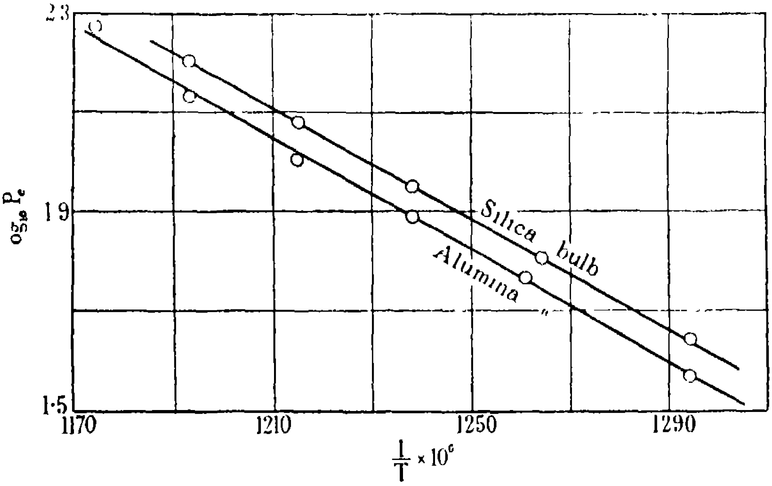


FIG. 3.—Variation of P_e with temperature according to an Arrhenius equation.

For this purpose we shall adopt the simple theory of the limit* which was worked out for the reaction between phosphine and oxygen and subsequently shown to apply in an approximate way to the hydrogen-oxygen system. No assumption about the chemical nature of the chain carriers is necessary, except the rather general one that two species X_0 and X_H are alternately produced and destroyed, and that there is a finite probability that the chain branches when X_0 meets H_2 or when X_H meets O_2. Thus X_0 or X_H is the X of the previous section while the Y of the previous section is now the H_2 or the O_2 of the theory of the lower limit. The derivation previously given† applies to the present system if PH_3 is replaced throughout by H_2 and X_P by X_H. The condition for explosion at the lower limit is found to be

---

* Dalton and Hinshelwood, 'Proc. Roy. Soc.,' A, vol. 125, p. 294 (1929), based upon work of Semenov, 'Z. Physik,' vol. 46, p. 109 (1927).

† The paper of Dalton and Hinshelwood, *loc. cit.*, contains on pp. 304 and 305 the derivation applying *mutatis mutandis* to the present requirements.

$$\nu p_{O_2} . p_{H_2} . \beta^2 . d^2/\lambda_0^2 = \text{constant}$$

in the absence of inert gases.* λ_0 is the average mean free path at unit pressure, d is the diameter of the vessel, and β is the factor by which the diameter is multiplied to give the " effective diameter " when the chains are reflected from the surface a number of times. ν, as in the last section, is the probability of branching.

For a given vessel and composition of gas mixture, neglecting the small variation of λ_0 with temperature, we have

$$p_{O_2} . p_{H_2} = \text{const.}/\nu\beta^2,$$

therefore

$$p_e = p_{O_2} + p_{H_2} = \text{const.}/\nu^{\frac{1}{2}}\beta,$$

whence

$$\frac{d \log p_e}{d\mathrm{T}} = -\frac{1}{2}\frac{d \log \nu}{d\mathrm{T}} - \frac{d \log \beta}{d\mathrm{T}}.$$

Even if β were constant, $\log p_e$ would decrease with temperature only half as fast as $\log P_e$ increases. β, the reflexion factor, will normally decrease as the temperature rises, owing to the greater tendency of the chain carriers to be destroyed in surface reactions. This factor still further diminishes the absolute magnitude of $d \log p_e/d\mathrm{T}$, and, being probably quite a considerable one, can easily reduce it to the almost zero value actually found.

Conclusion.

We are inclined to think that the general relationships governing the upper and lower limits are now fairly satisfactorily explained. The influence of hydrogen-oxygen ratio, inert gases, vessel diameter and temperature can be qualitatively and approximately quantitatively accounted for. In several respects the effects at the two limits are in sharp contrast as shown in Table VI.

Thus quite apart from any quantitative agreements there is a great enough variety of qualitatively significant material to place the simple theory on a very strong foundation.

It will be noted that the original hypothesis that the gas phase deactivation at the upper limit consisted in the mutual destruction of hydrogen peroxide molecules is now replaced by the more general one of ternary collisions in which any third molecule may participate.

---

* p_e refers to the *lower* limit; P_e to the *upper*.

Another respect in which the original hypothesis may now be simplified is the following. From the fact that the explosion is stopped by gas phase deactivation, while the chains, by which the reaction occurring at still higher pressures is propagated, are stopped chiefly at the wall, Thompson and Hinshelwood originally suggested that the explosion and the measurable reaction depended on two different kinds of chain. This complication may not now be necessary. In the explosive region the complex XY is supposed to be able to break up giving more than one entity which can continue the chain; if a third molecule is present the branching process becomes impossible, but there is no reason why a single chain carrier should not emerge even from the ternary collision, and propagate unbranching, non-explosive chains. For example, in the absence of a third molecule there might be a finite probability of the change XY = 2A + B, while in a ternary collision this probability is zero compared with that of the reaction XY = A_2 + B. If A and B can propagate chains, while A_2 cannot, then we have the state of affairs demanded. The normal unbranching chains become longer and more numerous as the pressure increases still further above the upper limit, as experiment has shown, and terminate principally on the wall[*] or by reaction with molecules such as iodine.[†]

Table VI.

	Lower limit.	Upper limit.
$H_2 : O_2$ ratio	Approximation to ideal hyperbolic relation between p_{H_2} and p_{O_2}.	Linear relation between partial pressures
Inert gases	Favour explosion. Argon more effective than helium	Tend to stop explosion. Argon less effective than helium.
Diameter	Approximation to $1/d^2$ relation.	No influence (Table III).
Temperature	Very small effect.	Considerable effect.

The theory of the explosion limits developed in this and previous papers does not depend upon the exact assumption one might make about the nature of the entities which we have represented by the symbols X_0, X_H and Y. We consider that this is a definite advantage, because there is still no certainty about the exact chemical nature of the reactions. We hope that further work may make it possible to identify the chain carriers with definite species such as H, OH, H_2O_2, and so on. This, however, must be done in a way which conforms to the generalized physical theory.

---

[*] Hinshelwood and Thompson, 'Proc. Roy. Soc.,' A, vol. 118, p. 170 (1928); Hinshelwood and Gibson, 'Proc. Roy. Soc.,' A, vol. 119, p. 591 (1928).

[†] Hinshelwood and Garstang, 'Z. phys. Chem.' (Bodenstein Festband), p. 656 (1931).

We are indebted to the Royal Society and to Imperial Chemical Industries, Limited, for grants.

Summary.

The influence of hydrogen-oxygen ratio, inert gases, temperature, and surface and diameter of vessel on the upper pressure limit of the " low pressure " explosion of hydrogen and oxygen has been investigated. The results can be explained in terms of the theory that at the upper limit a branching chain process is balanced by gas phase deactivation in ternary collisions. On this basis the variation of the explosion pressure with hydrogen-oxygen ratio and with the pressure of inert gases present can be quantitatively accounted for. The mass of the two chain carriers concerned in the ternary collision must be considerably greater than that of two hydrogen atoms.

The differing influence of temperature on the upper and lower limits is explained by the theory.

The alternative views that the upper limit depends upon surface adsorption relationships or upon various thermal conductivities are inconsistent with the experimental observations.

The Thermal Decomposition of Acetaldehyde and the Existence of Different Activated States.

By C. J. M. FLETCHER and C. N. HINSHELWOOD, F.R.S.

(Received March 27, 1933.)

1. *Introduction.*

Homogeneous thermal gas reactions were at one time tacitly assumed to possess a definite order, unimolecular and bimolecular reactions, for example, being sharply distinguished. The kinetics of the decomposition of acetaldehyde, $CH_3CHO = CH_4 + CO$, over the pressure range of 100 to 400 mm. were found to satisfy the criterion of a bimolecular reaction, namely, that the reciprocal of the time for half change ($1/t_{\frac{1}{2}}$) plotted against the initial pressure (p_0) gave a straight line inclined to the axes The line, however, did not pass through the origin, as may be seen in fig. 1 of the present paper. This indicated the presence of some first order reaction, the nature of which was not determined.*

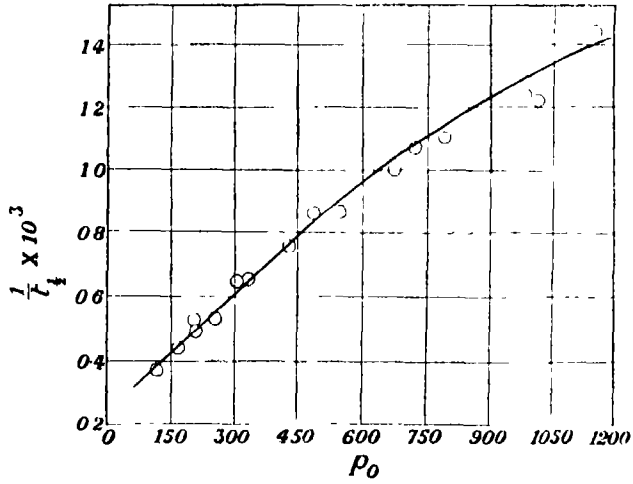


FIG. 1.

Subsequently, in accordance with the collision theory of activation and deactivation, it was shown that certain reactions, sometimes called quasi-unimolecular, change their order from the second at low pressures to the first at high pressures. This apparently was the reverse of the behaviour shown by acetaldehyde.

---

* Hinshelwood and Hutchison, 'Proc. Roy. Soc.,' A, vol. 111, p. 380 (1926).

The decomposition of nitrous oxide resembles that of acetaldehyde in that, although predominantly of the second order in the range 100–1000 mm., the straight line obtained by plotting $1/t_{\frac{1}{2}}$ against p_0 makes a positive intercept on the axis of $1/t_{\frac{1}{2}}$, when extrapolated to zero pressure.* Measurements at low pressures, however, do not correspond to this extrapolated straight line, the actual values of $1/t_{\frac{1}{2}}$ lying on a second line directed approximately toward the origin, and bending round rather sharply to the first line at a pressure of about 40 mm.† Thus the "first order" reaction is of the quasi-unimolecular type. As far as this evidence goes the decomposition of nitrous oxide would consist of a quasi-unimolecular reaction becoming of the first order above about 40 mm. and a simultaneous bimolecular reaction. But according to Volmer and Nagasako‡ this bimolecular reaction on investigation at higher pressures (5 to 10 atmospheres) itself proves to be quasi-unimolecular.

In view of this complexity of behaviour it is evident that the rather summary type of investigation originally adapted to decide between the simple alternatives of a unimolecular or bimolecular reaction is inadequate. We have therefore examined the decomposition of acetaldehyde over a much wider pressure range. When the results are plotted in the manner described above three distinct bends in the line are obtained, giving four segments of different slope. The simplest way to explain the curve is to assume that a number of independent quasi-unimolecular reactions occur, each of which passes from the second order to the first in a different range of pressure. The result, taken in conjunction with what is known about the nitrous oxide decomposition, suggests that such a form of curve may be general for molecules which contain more than two atoms but are not really very complex in structure. It has an interesting bearing upon the problem of the distribution of activation energy among different types of molecular vibration. This matter is discussed in section 3.

2. *Experimental Results.*

The experimental method was that described in a number of previous papers, the reaction being followed by pressure measurements. For the pressure range 15–1000 mm. a mercury manometer was used. The volume of the silica reaction vessel was about 200 c.c., and the "dead-space" in the apparatus amounted to less than 2 per cent.

---

* Hinshelwood and Burk, 'Proc. Roy. Soc.,' A, vol. 106, p. 284 (1924); Hibben, 'J. Amer. Chem. Soc.,' vol. 50, p. 940 (1928).

† Musgrave and Hinshelwood, 'Proc. Roy. Soc.' A, vol. 135, p. 23 (1932).

‡ Volmer and Nagasako, 'Z. phys. Chem.,' B, vol. 10, p. 414 (1930).

For the work at lower pressures three McLeod gauges were used, reading pressures correct to 1 per cent. in the ranges 40–4 mm., 10–1 mm., and 1–0·1 mm. respectively. As the use of these required the pressure in the reaction vessel to be shared with the evacuated gauge, each point on the decomposition-time curve for a given initial pressure needed a separate experiment. The aldehyde vapour was allowed to react for a known time, the final pressure being determined by sharing with the gauge. Initial pressures could be reproduced from experiment to experiment by keeping the aldehyde supply in ice, and using parts of the connecting tubes as measuring pipettes. The "sharing factors" for converting pressures in the gauge to pressures in the reaction vessel were carefully determined. They varied with the temperature of the vessel, the nature of the gas shared, the time during which the connecting tap was left open and with the actual magnitude of the pressure shared. By determining the sharing ratio with air and with acetaldehyde at temperatures where its rate of reaction was negligible it was found that the value for aldehyde was slightly lower. Allowance was made for this, and the influence of the other factors was then found by suitable experiments with air. The time required for the establishment of equilibrium through the capillary tubing and the tap was greater at lower pressures, but was always small in comparison with the half-time of the reaction, the maximum not exceeding 5 per cent. of this time. For pressures greater than 1 mm. the sharing ratio was independent of the pressure shared. Below 1 mm. it fell slightly, *e.g*, at 560° from 2·65 at 1·0 mm. to 2·62 at 0·3 mm. With this empirical knowledge of the minor variations in the sharing ratio the pressure in the reaction vessel could be determined with some accuracy in any of the pressure ranges.

Satisfactory reaction-time curves could be constructed from the various series of experiments, as may be seen from the examples shown in fig 2.

The calibration of the McLeod gauges was carried out by comparing the least sensitive gauge with the manometer directly, and then comparing the gauges among themselves at suitable pressures. The correctness of the calibrations was confirmed by the fact that rate measurements at certain pressures made independently with manometer and with McLeod gauge agreed.

The acetaldehyde was carefully fractionated (21°–22°) from the purest obtainable. It was stored in a glass bulb sealed to the apparatus, and lest it should deteriorate was replaced at frequent intervals.

The decomposition was followed at three temperatures, 560°, 516° and 496° C. The temperature of the furnace was controlled in many of the experiments by a Cooke and Swallow regulator* which kept it constant to within 1°. It was measured by a platinum-platinum-rhodium thermocouple.

The end-point of the reaction at moderate pressures in an unpacked vessel corresponded to a pressure increase of 98 per cent. of the theoretical. The products are known to consist solely of carbon monoxide and methane. At low pressures the end-point fell slightly, but down to 0·2 mm. remained above 90 per cent. Below 0·2 mm. initial pressure the measurements became uncertain, as the end-point was vitiated by adsorption of the aldehyde and of the reaction products to varying extents.

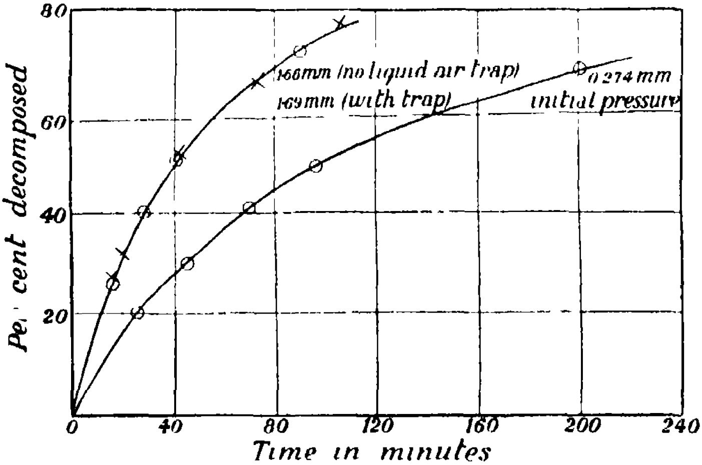


FIG. 2.— O without liquid air trap ; + with liquid air trap.

To find out whether mercury vapour influenced the results at low pressures in any way the manometer was removed altogether and a liquid air trap was inserted between the mercury vapour pump and the rest of the apparatus. The decomposition was measured by the McLeod gauge, but no mercury vapour from this could possibly diffuse into the reaction vessel, since the aldehyde vapour did not pass through the gauge on the way to the vessel and the sharing did not occur until after the reaction. Experiments at 6 mm. and at 1·7 mm. initial pressure were quite uninfluenced by the presence or absence of the liquid air trap. An example is shown in fig. 2.

To obviate any catalytic influence of traces of air the apparatus was always very thoroughly evacuated—to less than 10^{-3} mm. No measurable pressure increase occurred when the evacuated apparatus was left overnight. If there had been occasion to admit air, the reaction vessel was always washed out several times with acetaldehyde vapour at the reaction temperature before further use.

---

* 'J. Sci. Inst.,' vol. 6, p. 287 (1929). We are greatly indebted to the authors for imbling and adjusting a regulator for us.

Course of the Reaction.—As found by Hinshelwood and Hutchison, the reaction-time curve for any given initial pressure is nearly bimolecular. This fact appears in the tabulated results of the present work. For a second order reaction $t_{\frac{1}{3}} : t_{\frac{1}{2}} = 1 : 2$, and $t_{\frac{1}{2}} : t_{\frac{2}{3}} = 1 : 2$. Values of these ratios are tabulated for a number of different initial pressures. At lower pressures $t_{\frac{1}{2}}$ tends to become slightly less than double $t_{\frac{1}{3}}$.

In all the tables the initial pressure, p_0, is given in millimetres, the initial rate, $(dp/dt)_0$ in millimetres per minute, $t_{\frac{1}{2}}$ is given in seconds.

The initial rates, being obtained from tangents to the curve at the origin, are naturally less accurate than the values of $t_{\frac{1}{2}}$.

Influence of the Initial Pressure.—The three ranges 100–1000 mm., 10–100 mm., and below 10 mm. may conveniently be considered separately. Table I gives the results for the higher range at 496°. In fig. 1 $1/t_{\frac{1}{2}}$ is plotted against p_0. The points to be noted are that a linear relation holds up to about 500 mm., but that the extrapolated line does not pass through the origin, and further that at higher pressures the slope of the curve becomes smaller. Kassel* expressed the variation of rate with initial pressure by a 5/3 power law. If $\log (1/t_{\frac{1}{2}})$ is plotted against $\log p_0$ for our results over a certain range, a similar relation appears. But we do not consider this to possess any theoretical significance.

Table I.—Temperature 496° C.

p_0.	$(dp/dt)_0$.	$t_{1/2}$.	$t_{1/2}/t_{1/3}$.	Remarks.
1156	79	690	1·93	
1016	68	820	2 02	
983	67	805	1·99	
795	48	906	2·05	
726	44	930	2·00	
678	36	1000	2·02	
547	27	1150	2 06	
486	23	1160	2·02	
428	19·5	1315	2·12	
332	15·2	1500	2·08	
308	12·0	1525	—	
258	8·0	1875	—	
208·5	5·4	2030	—	(end-point 99 per cent.)
205	6·1	1890	2·08	
171·5	4·5	2275	—	
120	2·6	2680	2·06	

---

* 'J. Phys. Chem.,' vol. 34, p. 1166 (1930).

Table II —Temperature 516° C.

p_0.	$(dp/dt)_0$.	$t_{1/2}$.	$t_{3/4}/t_{1/2}$.	Remarks.
356	31	577	1 95	Manometer.
315	27	604	2 05	
263	22	660	2 08	
240	20	705	2 04	
187	14·7	820	1 95	
171	10 5	895	1 94	
119	6 6	1110	2 02	
79	3·4	1360	1 98	
54	1 8	1700	2 06	
41	1 3	1920	—	
25 8	0 6	2440	1 94	
9 4	0 12	4860	—	
41	-	1870	1·89	McLeod gauges I and II.
30·5	—	2340	1·98	
21	—	2800	2 04	
18 4	—	3270	1 97	
6 73	—	5640	2 00	
3 44	—	8400	2 06	

Table III gives results for 560° from 0 2 mm. to 400 mm. In fig. 3, $1/t_{\frac{1}{2}}$ is plotted against p_0 from about 3 to 200 mm., and inset is given the curve from 10 to 400 mm. Fig. 3 shows that the extrapolation of the high pressure line to an intercept on the $1/t_{\frac{1}{2}}$ axis is not actually followed. The line bends round

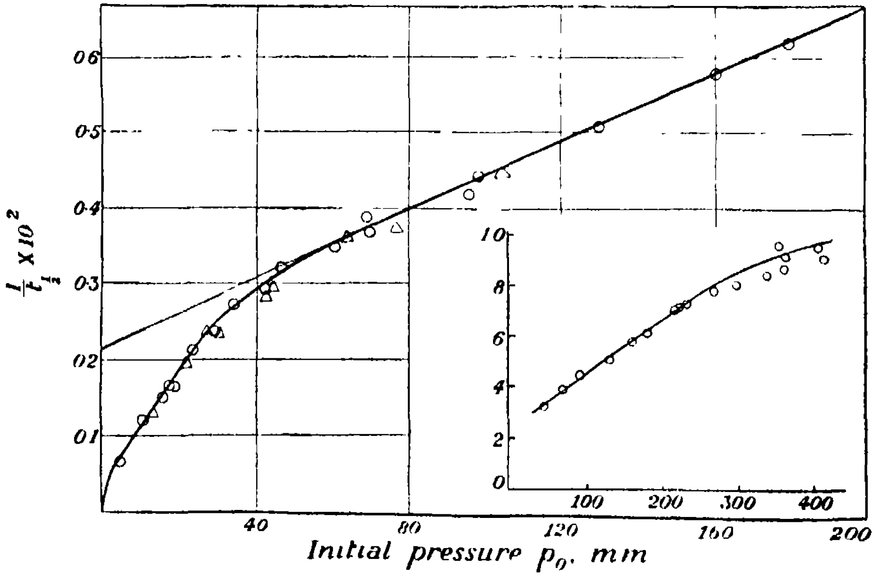


Fig. 3.— ○ unpacked bulb ; △ packed bulb, Temperature 560° C.

Table III.—Temperature 560° C. Unpacked bulb.

p_0.	$(dp/dt)_0$.	$t_{1/2}$.	$t_{1/2}/t_{1/3}$.	End-point.	Remarks.
				per cent	
415	200	111	1·93	—	Experiments with manometer
407	200	105	1·91	—	
362	170	110	1·87	—	
360	150	115	2·00	—	
355	150	104	1·93	—	
338	130	120	1·91	—	
296	120	124	1·91	—	
268	100	127	1·96	—	
233	80	137	1·98	—	
224	75	140	1·89	—	
216	75	142	1·97	—	
180	50	161	1·94	—	
161	40	173	1·92	—	
130	32	196	1·96	—	
91·5	20	225	1·92	—	
69·5	13·2	258	1·97	99	
46	7·5	310	1·94	—	
99	20·4	225	1·98	—	Experiments with manometer
94	18·0	238	1·94	—	
70·5	14·0	270	2·01	—	
60·5	10·2	285	1·87	—	
42	6·0	343	1·92	—	
33·6	4·8	365	1·87	—	
29·0	3·0	420	1·87	96	
23·2	2·6	470	1·96	—	
18·1	1·6	600	—	—	
17·0	1·4	600	1·87	—	
15·4	1·1	660	1·94	—	
10·0	0·5	840	2·00	96	
3·92	—	1530	1·93	95	McLeod II.
2·16	—	2160	1·96	—	
1·66	—	2400	1·86	—	$t_{2/3}/t_{1/2} = 1·88$
1·04	—	3120	2·08	—	McLeod III
0·576	—	4260	1·88	95	
0·274	—	6000	1·93	91	$t_{2/3}/t_{1/2} = 1·90$
					Liquid air trap
1·69	—	2400	1·90	—	$t_{2/3}/t_{1/2} = 1·89$
6·06	—	1116	2·04	98	$t_{2/3}/t_{1/2} = 2·04$

rather rapidly at about 40 mm. to a steeper line. Fig. 4, in which the low pressure range is drawn on an enlarged scale shows that the steeper line in fig. 3 is itself directed towards an intercept, but that another bend occurs before this is reached. Whether after this last bend the curve is directed exactly towards the origin is hard to say, though it may well be. Experiments in a still lower pressure range cannot be made for the reasons already explained.

Thus we have for the relation between $1/t_{\frac{1}{2}}$ and p_0 at 560° a line showing rather abrupt changes of direction at about 1 mm., at about 40 mm. and again in the neighbourhood of several hundred millimetres (the latter change corresponding to that shown in fig. 1 and noticeable in the inset of fig. 3*).

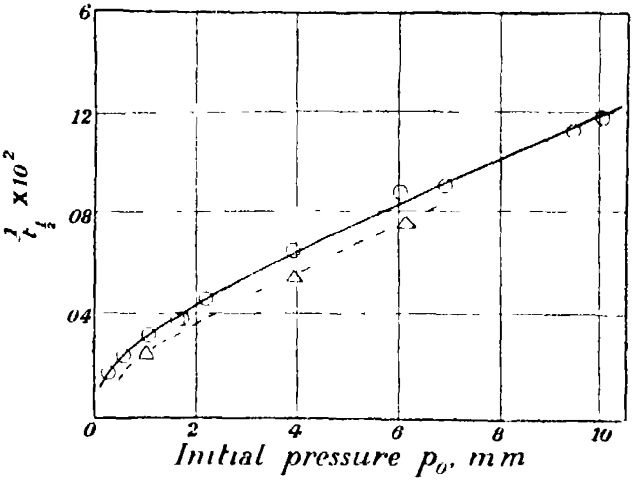


FIG. 4.— O unpacked bulb ; Δ packed bulb.

The separation of the curve into these segments is confirmed by experiments made at 516°. The results are given in Table II. In fig. 5 those for the range

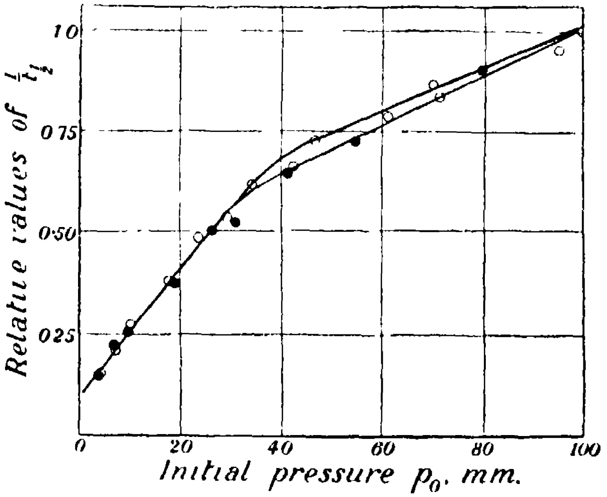


FIG. 5.— O = 560° ; ● = 516°.

---

* The products of reaction are not responsible in any way for the form of the curve, since the composite nature of the reaction is quite clearly revealed if the initial rate is plotted instead of $1/t_{\frac{1}{2}}$. A suitable method of representing the results is to plot $p_0{}^2$/initial rate against p_0, which would give a straight line for a simple quasi-unimolecular reaction, but does in fact give a line with abrupt changes of slope. The only advantage of using $t_{\frac{1}{2}}$ is that it can be measured much more accurately than an initial rate.

0–100 mm. are plotted. The values of $1/t_{\frac{1}{2}}$ are given as a fraction of the value at 100 mm., and the corresponding values for 560° are shown on the same scale so that the relative positions of the bend at the two temperatures can be seen.

Influence of Surface.—In Table IV are recorded the results obtained using a silica vessel packed with some silica tubes. The ratio surface : volume of this vessel was approximately four times as great as that of the unpacked vessel used in the other experiments. If the surface is increased a large number of times, as for example by introducing powdered silica, there may be an appreciable acceleration of the reaction, but this is accompanied by a serious disturbance of the end-point, owing to adsorption, polymerization of the aldehyde, or to both causes. The object being not to detect the existence of a few per cent. of surface reaction but to find out whether it was enough to account for the "first order" portion of the total change in any of the pressure ranges, the more moderate increase of surface provided by the tube-packed bulb was considered better than any drastic change. Even with the fourfold change the end-point was somewhat affected at the lowest pressures ; $t_{\frac{1}{2}}$ was always taken as the time corresponding to half the actually observed end-point. If the observed end-point is used the reaction-time curve is approximately bimolecular. Under these conditions the reaction in the packed bulb appeared to be actually rather slower.

Table IV.—Temperature 564° C. Packed bulb.

p_0.	$t_{1/2}$.	$t_{1/2}/t$.	End-point.
			per cent
144	190	—	—
105	222	2·02	-
76	268	2·09	
64	270	1·93	-
44	340	1·93	—
42	350	2·06	-
29·5	424	1·98	—
28	415	2·04	100
21·5	510	2·04	—
13	780	2·05	> 90
6·1	1410	1·88	80
3·96	1800	2·00	75
1·08	3900	1·97	61

In order that the shape of the curves for the packed and unpacked bulbs might be compared the temperature was raised about 4° for experiments with the former, so that the results at moderate pressures (100 mm.) coincided. From the disposition of the points for the packed bulb in figs. 3 and 4 it appears that the changes in slope are not due to changes in the relative proportion of surface and homogeneous reaction.

Influence of Temperature.—The rate of reaction was measured for initial pressure of 25 mm., 200 mm. and 450 mm. respectively at a series of temperatures (five for each pressure). log $t_{\frac{1}{2}}$ plotted against 1/T gives a straight line for each pressure. The values of E, the heat of activation, uncorrected for any variation of collision rate with temperature, are as follows :—

25 mm., E = 55,000 cal.
200 mm., E = 50,400 cal.; repetition, E = 49,700.
450 mm., E = 47,700 cal.

The uncorrected value found by Hinshelwood and Hutchison for initial pressures in the region of 400 mm. was 46,500 cal. approximately. This value is subject to slight error since the variation of E with pressure was not realized when the measurements were made. It should be mentioned that in the present work the measurements at 25 mm. and at 200 mm. were made in alternation on the same day, and those at 450 mm. made alternately on another day with a complete new series of measurements at 200 mm. This ensured that any absolute errors in the temperature should not affect the relative values of the rates. Thus it appears that the reaction which predominates at the higher pressures has a smaller energy of activation than that at lower pressures. The detailed results are given in Table V.

Table V.

Temperature ° C.	$\mathrm{Log}_{10}\,(10^3 \times 1/t)_{1/2}$.	
	25 mm.	200 mm.
615	(1·194)	—
601	1·049	—
581	0·732	(1·204)
560	0·377	0·832
540	0·170	0·509
521	—	0·170
502	—	$\bar{1}$·848

In plotting these little weight was given to the points for the highest temperature in each series.

Temperature ° C.	200 mm.	450 mm.
572·5	1·065	—
549·5	0·697	0·865
533	0·456	0·590
516	0·133	0·310
501	$\bar{1}$·883	0·062
488·5	—	$\bar{1}$·835

3. *Theoretical Discussion.*

We consider that the changes in slope of the curves showing $1/t_{\frac{1}{2}}$ as a function of p_0 are too abrupt to be explained except by the assumption that the reaction is kinetically composite. It seems to us that the simplest explanation lies in the existence of different modes of activation corresponding to particular divisions of the energy of the molecule among a limited number of vibrational (or rotational) states. The interchange of energy among these different modes is sufficiently difficult to give rise to a number of virtually independent quasi-unimolecular decompositions. The following simple considerations show how this assumption will explain the results.

Activation occurs by collision. Let the concentration, measured as a partial pressure, of molecules in a particular type of activated state be a. They are formed at a rate $k_1 p^2$, where p is the partial pressure of normal molecules, and are removed by deactivation at a rate $k_2 p \cdot a$, and by chemical transformation at a rate $k_3 a$. Thus $k_1 \cdot p^2 - k_2 \cdot p \cdot a - k_3 \cdot a = 0$. The rate of reaction $k_3 \cdot a = k_1 \cdot p^2/(1 + k_2 \cdot p/k_3) = -dp/dt$. From this we find

$$\frac{1}{t_{\frac{1}{2}}} = \frac{1}{1/(k_1 \cdot p_0) + (k_2/k_1 \cdot k_3) \cdot \log 2}.$$

When p_0 is small this gives $1/t_{\frac{1}{2}} = k_1 \cdot p_0$, and when p_0 is large

$$1/t_{\frac{1}{2}} = k_1 \cdot k_3/k_2 \log 2.$$

The complete curve is of the type 1, 2, and 3 of fig. 6. The height of the horizontal part is proportional to k_3, the probability of transformation of the activated molecules. (This probability is the reciprocal of the average life the activated molecule would possess if they were never deactivated by collision.) The bend in the curve occurs where $(k_2/k_3)p_0$ is comparable with unity, *i.e.*, the greater k_3 the higher the pressure. When k_3 is infinite, *i.e.*, the molecules react in the activating collision itself, the curve does not bend over at all (true bimolecular reaction).

If now we have a small finite number of modes of activation each associated with a different probability of transformation, the rate will be

$$-\frac{dp}{dt} = \Sigma \frac{k_1 . p^2}{1 + (k_2/k_3)\, p}. \qquad (1)$$

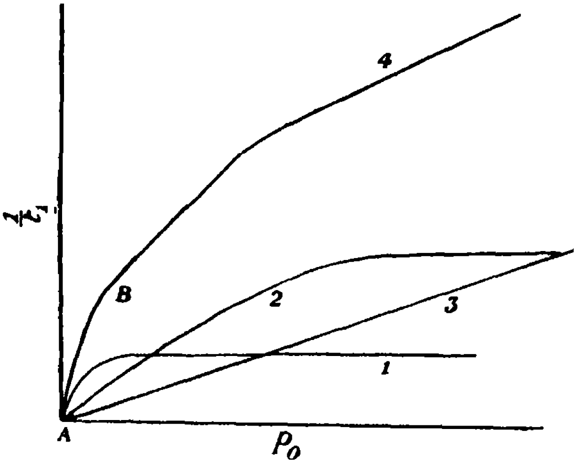


Fig. 6.

Suppose there were three such types of activated state, for which

$$k_3 \gg k'_3 \gg k''_3\,;$$

then at the lowest pressures the equation reduces to

$$-\frac{dp}{dt} = (k_1 + k'_1 + k''_1)\, p^2,$$

whence

$$1/t_{\frac{1}{2}} = (k_1 + k'_1 + k''_1)\, p_0.$$

This corresponds to the segment AB of curve 4 in fig. 6. There will be a range of pressure where the last term in (1) has reached its limiting value

$$(k''_1 . k''_3/k''_2)\, p,$$

which itself is small while the first two terms are still representable by k_1p^2 and k'_1p^2. The line for $1/t_{\frac{1}{2}}$ against p_0 will have a slope approximately equal to $(k_1 + k'_1)$ and be directed towards an intercept on the axis equal to

$$k''_1\, k''_3/k''_2 \log 2.$$

Similarly in another higher pressure range the slope will approximate to k_1

and the line will be directed towards a larger intercept. The result is a segmented curve as shown, with fairly abrupt changes of direction.*

Elaborations of the collision theory of activation have been worked out by Rice, Ramsperger, Kassel and others. These theories take into account the variation of the transformation probability with the magnitude of the energy of activation. Thus k_3 ceases to be a simple constant, but it will possess a definite average value. The present results, and those for nitrous oxide suggest that for molecules of a moderate degree of complexity there are a limited number of rather widely different k_3 values, associated with what are virtually independent decomposition mechanisms. This conclusion is not affected by the possibility that each k_3 is really an average value only.

One explanation is that the activation energy can be stored in several different modes of motion, *e.g.*, bond vibrations, and that there is a widely different probability of transformation according to those in which it resides. If the energy does not pass readily from one linkage to another in the absence of collisions then there will be independent quasi-unimolecular reactions.

If the k_3 values are not simple constants the actual form of the reaction-rate, pressure curve may not be expressible by an equation in any convenient way though its composite nature will be clearly visible. Certain types of collision might give immediate reaction, so that in addition to the quasi-unimolecular reactions there may be some change remaining of the second order up to the highest pressures.

The different activated states might be the successive quantum levels of a given vibration. One would then expect the greater k_3 value to be associated with the greater E value, *i.e.*, the mode of reaction predominating at the higher pressure would have the greater energy of activation. The results described in the section on the temperature coefficient indicate that the heat of activation decreases at higher pressures. Thus we are inclined to believe that the activated states are qualitatively rather than merely quantitatively different.

---

* For a quasi-unimolecular reaction with a single constant k_3, plotting $1/p_0$ against $t_{\frac{1}{2}}$ gives a straight line making an intercept on the axis equal to the value of $t_{\frac{1}{2}}$ at high pressures. This method of plotting is suitable for extrapolating to find the limiting value at high pressures, but is less satisfactory for exploring the low pressure region, on account of the distortion of the scale when the reciprocal of p_0 is taken. If the results in fig. 3 are plotted in this way the bend (at 30–60 mm.) is smoothed out into a fairly straight line from about 15 to about 80 mm., but for values of $1/p_0$ less than about 0 01 the line bends sharply down towards the axis. (This corresponds to the fact that the line in fig. 3 does not become horizontal above 100 mm.) Also for values of $1/p_0$ greater than about 0·07 the line curves markedly away from the $t_{\frac{1}{2}}$ axis, but the expansion of the scale of reciprocal pressures here tends to mask the importance of the change of slope.

We have investigated numerous methods of plotting the results and concluded that the one used in the text is the most appropriate for the present purpose. We are not in agreement with Volmer and Froelich (' Z. phys. Chem.,' B, vol. 19, p. 85 (1932)) if we are right in understanding them to consider that the behaviour found with nitrous oxide can be accounted for by a continuous variation of the probability of transformation, without assuming the reaction to be kinetically composite.

The following considerations are of an approximate nature only, and based upon the assumption that each k_1, k_2, and k_3 can be treated as a single constant. Since for a quasi-unimolecular reaction the curve for $1/t_{\frac{1}{2}}$ changes its slope rapidly where $k_2 . p/k_3$ is of the order of magnitude unity, the bends in the composite curve will occur where $k_2 . p = k_3$. If we can further assume that the different k_2 values are of the same order of magnitude, then, from the fact that successive bends occur in the ranges 0·5–2, 30–50, and again at several hundred millimetres, it will follow that three successive transformation probabilites increase in the ratio of one to two powers of ten.

Assuming that the lowest segment of the curve really passes through the origin, one can calculate from the intercepts and the position of the bends the following approximate values of k_1 at 560° (the units being seconds and millimetres); for the lowest pressure range $2{\cdot}4 \times 10^{-4}$, and for the next range $3{\cdot}3 \times 10^{-5}$. From the slopes the more certain value $5{\cdot}1 \times 10^{-5}$ is found for the latter. The sum of the k_1 values for all the higher ranges is $2{\cdot}3 \times 10^{-5}$. Thus k_1 values diminish for the higher pressure ranges, *i.e.*, those states which have the greater probability of decomposition are less frequently produced.

Nevertheless the corresponding energies of activation are smaller. There must therefore be some compensating factor. At the higher pressures the rate of reaction was previously found to be accounted for if the energy of activation was taken as 46,500 (uncorrected) and assumed to be distributed in two square terms only. The present values of E are somewhat higher, but diminish with increasing pressure. At the highest pressures E will probably not exceed 45,000 or 46,000, so that two square terms will still be enough. To account for the rate at the lower pressures the energy must be distributed over a larger number of degrees of freedom, as a simple calculation shows. That there should be a smaller probability of transformation when the energy is subdivided in the molecule is understandable, since redistribution must be waited for. This line of thought leads to the idea that the various kinds of activated molecule may differ not so much in the assignment of the greater part of the activation energy to one or other linkage, nor yet in the number of quanta in a particular link upon the breaking of which decomposition principally depends, but on the number of degrees of freedom from which the energy has to be collected into this link. For a molecule of fairly simple structure the number of possibilities is limited.

It would be easy to speculate at length on the possible connection between the number of types of activated state revealed by experiment and the number of bonds in the molecule (the two are at least of the same order of magnitude), or, on the other hand, on the identification of the different states with a localization of the principal part of the energy in specific single bonds. Such speculations can profitably await the accumulation of more experimental evidence. In the meantime it may be said that, apart from the question of detailed interpretation, experiment seems to show quite clearly that distinct types of activated molecule must be assumed to account for the kinetic behaviour of certain reactions.

We are indebted to the Royal Society and to Imperial Chemical Industries for grants.

Summary.

The rate of decomposition of acetaldehyde has been measured over a range of initial pressures 0·2 to 1100 mm.

The curve obtained by plotting the reciprocal of the time of half decomposition against the initial pressure shows abrupt changes of slope, which can only be explained by assuming the reaction to be kinetically composite, even when wholly homogeneous. The results can be interpreted by the theory that the acetaldehyde molecule can be activated in a limited number of distinct ways, and that the different activated states are associated with different transformation probabilities. There are thus several virtually independent quasi-unimolecular decompositions for the same chemical reaction.

The energy of activation for the mode of reaction predominating at low pressures is higher than for greater pressures.

The nature of the various types of activated state is discussed.

Notes on Some Electronic Properties of Conductors and Insulators.

By R. H. Fowler, F.R.S.

(Received April 13, 1933.)

§ 1. *Introduction. Tamm's Surface Levels.*—This paper contains a number of minor results to which I have been led in attempting to co-ordinate various phenomena exhibited by semi-conductors and insulators, and their contacts with metals. It is essentially a critical survey. Much of it has obviously been inspired by the beautiful theory of the "surface levels" proposed by Tamm,[†] which I have strenuously tried to incorporate significantly into the general theory. It is natural at first sight to experiment with the view that the existence of these surface levels in crystalline insulators is one fundamental reason for their insulating properties. This view, however, one is forced to abandon, as will be shown here. The insulating property is not after all so subtle. But I believe the discussion I shall present will help to clear up ideas on these subjects and on a number of allied ones, which have been until recently more obscure, at least to me, than they need have been.

After considering, therefore, the nature of insulation by a crystal (as opposed to an amorphous glass) we discuss it in terms of Tamm's surface levels and show what properties are required of these levels if they are to be responsible for the insulation. These properties are possible but peculiar. We then examine the origin of the surface levels more exactly and show how they have to fit into the scheme of lattice levels. It is not possible to conclude that they cannot be concerned in the insulation, but it does seem possible to conclude that it is unlikely that they are concerned in any fundamental way. We then take up the question of the surface conditions on an insulator on a wider basis and show that the necessary surface charge of extra or missing electrons can always be accommodated on the conductors in contact with the insulator which separates them. The paper concludes with certain applications of these ideas of surface conditions to the phenomena of photoconductivity. The main subject matter of the paper has close affinities with von Hippel's[‡] work on electrical breakdown in crystalline insulators but the same questions are not handled. None the less, it is important to record that everything we have to say is in conformity with his theory.

[†] 'Phys. Z. Sowjet Union,' vol. 1, p. 733 (1932).

[‡] 'Z. Physik,' vol. 67, p. 707; vol. 68, p. 309 (1931); vol. 75, p. 145 (1932).

§ 2. *How a Crystal Insulates.*—From the phenomena of photoconductivity, for example the photoconductivity of zinc blende or diamond, one must conclude that a perfect crystalline lattice can only act as an insulator because it contains no free electrons in any band of running levels which should be empty, and no " free holes " or vacancies among the electrons in any band of running levels which should properly be full. Any concentration of free electrons in any nearly empty band, or of free holes in any nearly full one, immediately endows the material with a conductivity which can be calculated according to the classical theory of Drude and Lorentz. It is perhaps not necessary to think of an amorphous glass in this way. Such a substance may possess no running levels through which an electron can move freely once it reaches them, although the conductivity of liquid mercury (which is presumably an " amorphous glass " from this point of view) is so similar to that of the crystalline solid near the melting point that it is unwise to dogmatize. Whatever may be the correct view for amorphous insulators, we have to admit that perfect crystalline insulators can exist, capable of standing up to large potential gradients, but that electrons or holes once they get inside, will move freely through the crystal. In order that a crystal should insulate, it is necessary therefore that no free electrons or free holes should be found in *or penetrate into* the running levels. In the ideal case near the absolute zero, this reduces to a condition at each surface of the insulator. The conductor-insulator interface must charge up (positively or negatively) to such a degree that the running levels are kept out of danger. It is in the problem of accommodating these surface charges that Tamm's surface levels (which belong to the crystal) might perhaps come into play. Before discussing this, however, we will complete the examination of how the running levels must arrange themselves relative to the levels in the metal conductors in contact with the insulator.

Fig. 1 shows approximately how the electron levels in the crystal insulator and the ideal metal electrodes must arrange themselves according to Wilson's theory, with the top of Sommerfeld's distribution in the metal in the middle of the relevant forbidden band of energies in the crystal. There is here no potential difference between the electrodes. Fig. 2 shows the state of affairs when a potential difference is established. Surface A must charge up negatively and surface B positively, so that (the potential being a continuous function) we get to the state of affairs in fig. 2. There being no mobile electrons anywhere in the crystal, there can be no space charge and the potential slope must be uniform there. We assume that the necessary electrons can be accommodated in and removed from "interface levels" without putting free electrons or free holes into the running levels.

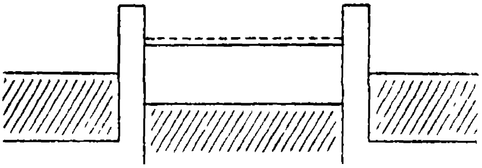


FIG. 1.—Electron levels in a crystal insulator in equilibrium between metal electrodes. The shaded bands represent energy levels full of electrons. The broken line indicates the presence of a band of empty energy levels through which electrons could move freely. The intervening band of levels in the crystal is forbidden.

In the state described by fig. 2, a small leak current will actually flow owing to the strong field effect. If the unit of current is amperes/cm.^{2}, energies are

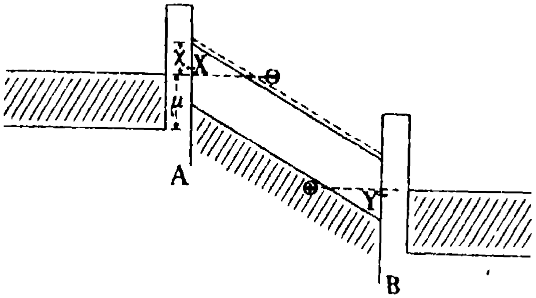


FIG. 2.

all expressed in electron volts, the gradient is F* volts/cm. and the effect of any permanent interface barrier is neglected, then the theory of Fowler and Nordheim† will apply and

$$I_- = I_+ = 6{\cdot}2 \times 10^{-6} \frac{\mu^{\frac{1}{2}}}{(\chi + \mu)\chi^{\frac{1}{2}}} F^{*2} e^{-6 \cdot 8 \times 10^7 \chi^{3/2}/F^*}. \tag{1}$$

where I_-, I_+ are the currents carried by free electrons and holes respectively. The parameters χ and μ are energies defining the effective work function of the metal against emission into the crystal (here the half height of the forbidden band) and the maximum kinetic energy of Sommerfeld's electrons in the metal respectively. If, for example, we take $\chi = 3$ volts, $\mu = 7$ volts

$$I_{\pm} = F^{*2}\, 10^{-6 - 1 \cdot 54 \times 10^8/F^*}, \tag{2}$$

---

† 'Proc. Roy. Soc.,' A, vol. 119, p. 173 (1928).

There is a similar formula for $\chi = 1$. These give leak currents as under :—

F^*	$\chi = 3$	$\chi = 1$
10^5	10^{-1500}	10^{-300}
10^6	10^{-147}	10^{-24}
10^7	10^{-7}	10^{-5}

Thus at low temperatures there can be almost perfect insulation up to fields as strong as 10^6 volts/cm. for any reasonable width of the forbidden energy band in the crystal, and perhaps up to strengths as great as 10^7 volts per centimetre for specially suitable lattices.

It is no good considering greater field strengths than 10^7 volts/cm. for at about this value, even in an ideal crystal, the few electrons that do enter will begin to acquire a kinetic energy of about 10 volts in a mean free path of 10^{-6} cm. Thus in this voltage neighbourhood at the latest electrical breakdown must begin to occur on account of cumulative " ionization " inside the crystal, which means the creation of free electrons and free holes. This, of course, is very crude and must be refined after the manner of von Hippel. It is easy to show that, for any current which can flow before cumulative ionization sets in, space charge effects can be neglected, and the potential gradient shown in fig. 2 is correct.

§ 3. *Surface Accommodation for the Electrons by Tamm's Levels.*—We pass on now to consider how the state of affairs required and shown in fig. 2 can be brought about. It is interesting to argue thus (and this argument is in effect implied by Tamm) : the interface charges may be thought of as held on the insulator surface. It is therefore necessary that at interface A there should be plenty of empty Tamm levels below the upper band of empty running levels, say at level X in fig. 2. These could then accommodate the electrons required to raise the electron energy levels at A. At the same time, it is necessary that there should be plenty of full Tamm levels above the full band of running levels, say at level Y. These could then supply electrons for removal to produce the positive charge necessary to depress the levels at B. Both sets are equally necessary to prevent entry of free electrons at A or entry of free holes at B, either of which would be equally fatal to the insulation.

It is possible that Tamm's levels might exist and fulfil both these conditions. To see this clearly, it is necessary to examine the origin of these levels somewhat more deeply than Tamm has done, using for this purpose a *finite* linear chain of N atoms, in regular spacing, each of which provides one orbital wave function capable of accommodating one or two electrons. When the spacing constant of this array is reduced, the atomic states of course fuse together in the well-known way into the lattice states, but in so doing the *number* of available states derived from the given set of N atomic states, cannot change. Tamm has shown that for a semi-infinite array there may be in any forbidden band, a surface state, which is not a periodic state running right through the array. But this state, even if it has special properties, cannot be an extra state and must be an outlying lattice state, one of the states into which the atomic states have fused. It is easy to overlook this point unless one considers a finite chain of atoms. In that case, generalizing Tamm's argument, the conditions for a special surface state can be satisfied at either end of the chain, so that if the given set of N atomic states gives rise to Tamm's surface states at all, it will give rise to them in pairs. Moreover, the lattice states arising from one set of atomic states can only give rise to Tamm's levels in the two forbidden bands immediately on either side of their own band of running levels. Otherwise the approximation which confines attention to one set of atomic wave functions breaks down. There are therefore *a priori* four alternatives. The set of N atomic states may give rise (1) to no surface levels, or (2) to upper surface levels above the running levels and no lower surface levels below the running levels, or (3) to lower levels and no upper levels, or (4) to both upper and lower levels. If Tamm's conclusions for the semi-infinite chain can be carried over unchanged to this finite chain, which is not quite certain, then alternatives (1) and (3) cannot occur, but either (2) or (4) may. A full band (2N electrons) will fill up all its running and surface levels, upper and lower alike. All the levels of an empty band must normally be empty.

In the conditions of fig. 2, the full upper Tamm levels of the full lower band will always be available at Y, as a source of supply of fixed positive charge. The empty lower Tamm levels of the upper empty band may or may not be available at X.

The Tamm levels, if present, could obviously do the work of accommodating the necessary interface charges closely localized to a plane.

§ 4. *Surface Accommodation by the Levels of the Metal.*—It is, however, quite unnecessary to cast the duty of accommodating these electrons on the surface levels of the crystal. The ordinary levels in the metal at the top of Sommerfeld's distribution will themselves act in exactly the necessary way. Since, however, we are going to add to or subtract from the electrons of a neutral metal, space charge effects will become important and must be brought into the theory. The first approximation can bring in the space charge effects, in pushing the extra free electrons (or holes) to the surface by using Thomas' method. This will give a sufficiently good approximation. Provided it turns out that these levels, when concentrated to the surface by the space charge, accommodate the extra electrons or the extra holes close to the level of the top of the undisturbed Sommerfeld levels, then they will for all our purposes function exactly like Tamm's levels at X or Y respectively, and Tamm's levels will be irrelevant to this particular problem.

Let V (x) be the electrostatic potential inside the metal at a distance x from the surface, which may be treated as plane, and ε the algebraic charge on one of the extra electrons (holes). The zero of V (x) may here conveniently be taken to be at the top of the undisturbed Sommerfeld levels infinitely deep in the metal. V satisfies Poisson's equation

$$\frac{d^2\mathrm{V}}{dr^2} = -4\pi\rho = -4\pi\varepsilon n,$$

where n is the volume concentration of the electrons (holes). We assume with Thomas that n may be determined by taking 2 electrons per h^3 of that phase space belonging to unit volume of metal at depth x in which the electrons have negative total energy. This condition on the kinetic energy is

$$\text{K E.} \leqslant -\varepsilon\mathrm{V} + (\text{K.E.})_0,$$

where $(\text{K.E.})_0$ is the kinetic energy at the top of Sommerfeld's distribution. Thus

$$v^2_{\max} = -2\varepsilon\mathrm{V}/m + v_0^2,$$

and the momentum $p_{\max}$ is given by

$$p_{\max} = mv_{\max} = m\,(v_0^2 - 2\varepsilon\mathrm{V}/m)^{\frac{1}{2}}.$$

The total extra volume of momentum space is therefore

$$\tfrac{4}{3}\pi m^3\,\{(v_0^2 - 2\varepsilon\mathrm{V}/m)^{\frac{3}{2}} - v_0^3\}.$$

It will prove that $|\varepsilon\mathrm{V}| \ll \frac{1}{2}mv_0^2$. Therefore the extra volume reduces sufficiently nearly to

$$-4\pi m^2 v_0 \varepsilon\mathrm{V},$$

and the extra electrons to $-8\pi m^2 v_0 \varepsilon\mathrm{V}/h^3$ per cm.^{3}. Poisson's equation is therefore

$$\frac{d^2\mathrm{V}}{dx^2} = \alpha^2\mathrm{V} \qquad \left(\alpha^2 = \frac{32\pi^2 m^2 v_0 \varepsilon^2}{h^3}\right). \tag{4}$$

The required solution of (2) is

$$V = Ae^{-\alpha x} \qquad (\alpha > 0),$$

in which A must be fixed by the total excess charge per cm.^{2} of surface. If n_σ is the surface density in electrons (holes) per cm.^{2}, then

$$n_\sigma = \int_0^\infty n dx = -\frac{1}{4\pi\varepsilon}\int_0^\infty \frac{d^2V}{dx^2}\, dx,$$

$$= -\frac{1}{4\pi\varepsilon}\alpha A. \tag{5}$$

Thus

$$V = -\frac{4\pi\varepsilon n_\sigma}{\alpha} e^{-\alpha x} \qquad \left(\alpha = \frac{4\pi m\, |\varepsilon|\, (2v_0)^{\frac{1}{2}}}{h^{\frac{3}{2}}}\right). \tag{6}$$

We may take v_0 for an ordinary metal to correspond to about 10 volts—the value is not important. Then $\alpha = 2 \times 10^8$ cm.$^{-1}$. The surface charge σ is related to the field strength by $F = 4\pi\sigma$. Thus if F^* is the field strength in volts per centimetre

$$\sigma = F^*/1200\pi,$$

and $n_\sigma = 5{\cdot}5 \times 10^5\, F^*$ electrons/cm.^{2}. We thus find

$$V = \pm\, 1{\cdot}7 \times 10^{-11}\, F^*\, e^{-2\times 10^8 x}.$$

A field of 10^8 volts/cm. is an absolute upper limit at the surface of a metal. For such a field the surface charge by spreading into the metal produces a potential difference between the inside and the surface of only 2×10^{-3} volts, and the surface effect dies away very rapidly, being insensible a few Angstroms from the surface, that is, within two or three layers of atoms.

From the point of view of providing surface electrons for the surface charge, any possible surface charge, positive or negative, can therefore be accommodated in the levels of the metal, strongly concentrated to the surface and without any significant departure in energy level from the top of the Sommerfeld distribution. Such electron concentrations will do everything that is required of surface charges and allow a crystalline insulator to behave as described in section 2.

By every canon of physical theory, we are therefore compelled to conclude that Tamm's surface levels should not be considered in the problem of insulating crystals. At most they can play a trivial part. But this conclusion in no way implies that they may not be important in some of the other connections, such as triboelectrical phenomena, suggested by Tamm.

§ 5. *Conduction through an Insulator with a Small Natural Conductivity.*—It is of some interest to extend the considerations of section 2 to the case where in any volume of the crystal there is a supply of free electrons and holes, small but sufficient to give a conductivity which is large compared with the conductivity arising from the electrons and holes entering from the outside electrodes. This is another limiting case in which the passage of the current only slightly perturbs the normal distribution. In such a case when the field is applied, the internal free electrons and holes at once move, but at first no entry of new ones from outside is possible. The movement therefore builds up space charges near the boundaries which flatten the internal gradient and steepen the boundary gradients until the external electrodes can supply by the strong field effect just as many free electrons and free holes per second as the internal gradient can carry through the body of the crystal by its ordinary conductivity.

The (idealized) arrangement which must result when the steady state is reached is shown in fig. 3. Actually, of course, the space charges will not be localized in a geometrical plane, and the gradient will change from F'_0 to F_1 and F_1 to F''_0 continuously. If σ', σ'' are the conductivities of the crystal by free electrons and free holes respectively, then the problem is defined and soluble giving the total current I as a function of V, the total applied potential, via the following six equations :—

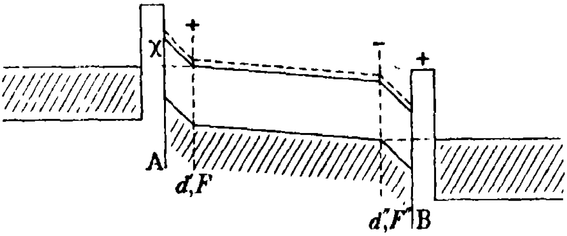


FIG. 3.

$$I(F'_0) = \sigma' F_1, \qquad I(F''_0) = \sigma'' F_1, \tag{7}$$

where I (F) is the function of F given by (1) or (2),

$$d' F'_0 = d'' F''_0 = \chi, \tag{8}$$

$$F_1 (d - d' - d'') + F'_0 d' + F''_0 d'' = V, \tag{9}$$

$$I = I(F'_0) + I(F''_0), \tag{10}$$

from which d', d'', F'_0, F''_0 and F_1 can be eliminated.

As the natural conductivity is diminished, the gradient F_1 must become steeper to carry the given currents until eventually it becomes equal to F'_0 and F''_0. But before this limit is reached, it is no longer possible to regard the current-carrying electrons as a small perturbation in the equilibrium distribution and the foregoing approximation breaks down.

§ 6. *Other Types of Boundary Conditions.*—In formulating the boundary conditions shown in fig. 3, section 5, we have ignored any potential hills which may exist on the interface—such hills are shown diagrammatically in the figure, but no account is taken of them in the calculations. It is, in effect, assumed that the free electrons all go through the right-hand boundary and the free holes all through the left-hand boundary for any reasonable adjustments of potential there. Such adjustments are then made as are required by the free electrons on the left-hand boundary and the free holes on the right-hand boundary.

This is an ideal limiting case. If there are potential hills of importance on the interfaces, the adjustment will be modified. On either boundary such a potential must build up that the combined movement of free electrons plus free holes across the barrier will carry the total current. At the same time the local space charges in the insulator must build up by adjusting the relative number of free electrons and free holes until the rates of dissociation and recombination thereby distorted strike separate balances between the incoming and outgoing free electrons and the incoming and outgoing free holes. There are just sufficient conditions to enable the requirements of the problem to be satisfied, but we shall not go into further details.

Similar considerations can be applied to insulators with weak conductivities which are due to free electrons alone, the holes being fixed and making no contribution to the conductivity. Again space charges must build up just inside the two boundaries sufficient to enable the boundary layer to transmit the proper number of free electrons. There will be a negative space charge of importance inside the insulator at the face where the electrons emerge only when there is a potential hill of some importance at the interface. At the other face where the electrons enter, there must always be a positive space charge to get the electrons in.

These *a priori* considerations as to the necessary nature of crystalline insulators can be satisfactorily related in a general way to the electrode phenomena of photoconducting crystals.† A photoconducting crystal is in general a fair insulator in the dark, to whose electrodes the foregoing theory will apply. Light produces free electrons and perhaps also free holes in those parts of the crystal to which it has access. They can then travel to other parts of the crystal, which may be in the dark, under the influence of an electric field. The effects described by Hughes and Du Bridge, in the paragraph quoted, can mostly be satisfactorily accounted for in terms of the theory here expounded. It is not necessary to repeat their descriptions, but one point must receive additional comment:—in order that electrons should enter or leave the crystal it is essential that the mobile charges should be able to reach the electrodes and there build up the necessary boundary fields Thus in particular, if a current is to continue to flow for a given voltage and illumination and not fade to zero, electrons must enter and leave the crystal and the boundary charges must be established

---

† See Hughes and Du Bridge, "Photoelectric Phenomena," McGraw Hill, 1932, chap. 2, especially § 8.5.

On the theory here developed, a continued current can always flow if the whole crystal is illuminated. This appears to agree with the facts. But if only a small portion of the crystal is illuminated not in contact with the electrodes, the final steady current is stated to be still non-zero for some crystals (*e.g.*, diamond, zincblende), but zero for others (*e g.*, yellow rock salt). In cold yellow rock salt the positive charges are isolated sodium atoms† and certainly *immobile.* Thus the necessary positive space charge can never be built up at the left-hand electrode of fig. 3, and electrons cannot enter, and the final current must be zero. In the other crystals on the other hand, such a positive charge must build up in what is claimed to be a dark region. This is possible so long as the positive charges are to some extent mobile (free holes). They are known to be sufficiently mobile in zincblende,‡ but are almost certainly immobile in diamond.§ If, therefore, any theory of this general form is to apply there is here a difficulty which needs further examination. It may be that it is difficult to ensure that the part of the diamond near the electrode from which electrons must enter is sufficiently in the dark.

An application of the same theory explains at once the great difference between the photoconduction by single crystals of selenium (red variety) and sulphur and that of zincblende, yellow rock salt, or diamond. In selenium and sulphur, practically no sign of saturation can be found for any ordinary field strengths for a given time of illumination, and the total charge flowing through the circuit may vastly exceed one electron per quantum of light absorbed. In the other substances the charge saturates as the voltage is raised at one electron per quantum (or less). This can at once be understood as an electrode effect. In selenium and sulphur we can suppose that sufficient space charges are immediately and easily built up near the electrodes (the charge required may be very small), so that the photoconductor will carry temporarily a current obeying Ohm's law. But in the other substances we must say that space charges are required near the electrodes which are not available for short illuminations and no new electrons can enter the photoconductor, except to neutralize a free hole. The best that can be done is to use all the electrons and holes set free by the light, which is the best that is observed.

---

† Wilson, 'Nature,' vol. 130, p. 913 (1932).

‡ Hughes and Du Bridge, *loc. cit.*, p. 297.

It is perhaps of interest to suggest that in view of the complications introduced by the electrodes, it might be of interest to attempt to develop a technique in which photoconductivity is measured in a circuit without electrodes, lying entirely within the body of the crystalline photoconductor. It might in this way prove possible to study effectively the steady (or time-variable) conductivity maintained in the crystal by steady illumination over a long period of time, freed from the complications introduced by space charges and electrodes. It seems from the tentative theoretical considerations of the succeeding sections that there is important information about the nature of crystalline insulators to be gained by such studies, at any rate in crystals where the secondary effects of the passage of current are not too complicated. It is not, however, suggested that this information will prove comparable with the information already obtained from the study of the transient phenomena of short-time illumination, a study which has been so successful in the hands of Gudden and Pohl.

§ 7. *Photoconductivity freed from Electrode and Space-Charge Effects.*—In this section we shall discuss the conductivity which would be exhibited by model crystals with certain specified properties if they were subjected to constant uniform illumination and the conductivity were measured free from effects of electrodes and of space charges. We shall also consider to some degree the rates at which the equilibrium value of the conductivity would be attained when the constant illumination is suddenly changed to a new value. We do not aim at accounting for the properties of any actual substance in particular because the conductivity has never been measured under conditions strictly comparable with those we shall have to assume. But in fact, the theory seems to reproduce in a general way some of the peculiar properties of the conductivity of selenium.

Consider a model crystal normally an insulator or intrinsic semi-conductor. The sensitivity to light is probably to be derived from the following primary process described by Wilson (*loc. cit.*). The light is absorbed by a molecule, NaCl, C_2 (in diamond), Se_2 (in selenium), producing a molecule in a new type of state, which acts as an impurity in the lattice. This impurity is probably Na + Cl *vice* Na^+Cl^- in rock salt, and the polar states C^+C^-, Se^+Se^- *vice* the homopolar states in the other examples. From these impurity states so formed, which in themselves give rise to no conductivity, conducting free electrons and (or) free holes are produced either by thermal effects or by further illumination as in rock salt. This conductivity may make use of all the electrons supplied by the impurities or only of a small temperature dependent fraction of them, according to the relative positions of the energy levels. All these points have been sufficiently discussed by Wilson (*loc cit.*) to show the type of effect to be expected. We shall discuss here the conductivity as a function of the illumination I when the illumination controls only the primary production of the impurities in the lattice.

The energy levels, the important transitions and their frequencies of occurrence, are shown in fig. 4. The dependence of the transition probabilities on the concentrations is shown *explicitly*. The coefficients ζ, λ, ρ, are functions of the temperature only (for a given crystal) ; κ might conceivably depend on the temperature, but is probably effectively a constant. The concentration n is the number of free conducting electrons per cubic centimetre, n_0 the number of free holes, so that the conductivity σ is $\alpha n + \beta n_0$; n^- is the concentration of impurity atoms ready to pass up an electron to the running levels (*i.e.*, the concentration of C^- or Na) and n^+ is the concentration of the other type of impurity (C^+ or Cl). The creation of an impurity pair does *not* produce holes in the full band of running levels. It merely diminishes the number of levels and the electrons in the full band by equal (insignificant) amounts. The + and – impurity levels are created and destroyed by recombination in pairs at rates $\kappa I + \rho_u$, $\rho_d n^- n^+$ respectively, and send electrons to and receive them from the upper levels at the rates $\lambda_u n^-$, $\lambda_d n n^+$. Besides

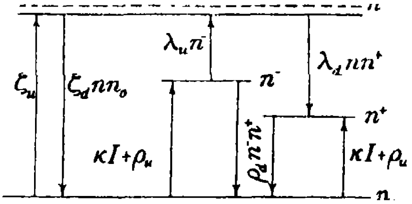


FIG 4.

these processes based on the illumination, there must be a connection of ordinary collision type between the upper and lower running levels, leading to the probabilities ζ_u, $\zeta_d n n_0$ for the upwards and downwards transition respectively. These processes do leave free holes in the lower levels. The equations governing the concentrations are then

$$\frac{dn}{dt} = \zeta_u - \zeta_d n n_0 + \lambda_u n^- - \lambda_d n n^+, \tag{11}$$

$$\frac{dn^-}{dt} = \kappa \mathrm{I} + \rho_u - \lambda_u n^- - \rho_d n^- n^+, \tag{12}$$

$$\frac{dn^+}{dt} = \kappa \mathrm{I} + \rho_u - \lambda_d n n^+ - \rho_d n^- n^+, \tag{13}$$

$$n = n_0 + n^+ - n^-. \tag{14}$$

The conductivity is

$$\sigma = \alpha n + \beta n_0, \tag{15}$$

in which it will often be accurate enough to omit the term βn_0 and take $\sigma \propto n$. These equations are valid only so long as all the n's are small compared with the number of electron levels in a band

The equilibrium state for constant illumination is given by the equations

$$\lambda_u n^- = \lambda_d n n^+, \tag{16}$$

$$\zeta_u = \zeta_d n n_0, \tag{17}$$

$$\kappa \mathrm{I} + \rho_u = \rho_d n^- n^+ + \lambda_u n^-, \tag{18}$$

$$n = n_0 + n^+ - n^-. \tag{19}$$

These lead to the quartic equation

$$\frac{\rho_d \lambda_u}{\lambda_d}\left(n^2 - \frac{\zeta_u}{\zeta_d}\right)^2 + \lambda_u n\left(\frac{\lambda_u}{\lambda_d} - n\right)\left(n^2 - \frac{\zeta_u}{\zeta_d}\right) - (\kappa \mathrm{I} + \rho_u)\, n\left(\frac{\lambda_u}{\lambda_d} - n\right)^2 = 0. \tag{20}$$

The other n's are all uniquely determinable in terms of the proper root of n from this equation.

A variety of behaviour is now possible. But the most natural type to expect is one in which ρ_u is very small, so that the "polar states" are only produced by light absorption. In that case when $\rho_u = 0$, $\mathrm{I} = 0$ then $n^2 = \xi_u/\xi_d$ and we find the ordinary conductivity of an intrinsic semiconductor $\sigma = \sigma_0\, e^{-\frac{1}{2}\mathrm{W}/k\mathrm{T}}$, where W is the energy gap between the two bands. It is easy to see the origin of the factor $e^{-\frac{1}{2}\mathrm{W}/k\mathrm{T}}$ in this presentation. The ζ's are collision probabilities and ζ_u will depend on T primarily though the Boltzmann factor $e^{-\mathrm{W}/k\mathrm{T}}$ since ζ_u is the probability that a collision between an electron in the lower band

and a lattice vibration will raise the electron to the upper band for which process the lattice vibration must have energy at least equal to W.

When I becomes fairly large, the last term in (20) will dominate, and we get

$$\mathrm{I} \rightarrow \infty, \quad n \sim \frac{\lambda_u}{\lambda_d} \; n_0 \sim \frac{\zeta_u}{\zeta_d} \frac{\lambda_d}{\lambda_u}. \tag{21}$$

The value of n saturates and σ saturates at a value in general much larger than the dark value The values are related by the equation

$$\frac{(\sigma_{\text{light}})_{\text{sat}}}{\sigma_{\text{dark}}} = \frac{\lambda_u/\lambda_d}{(\zeta_u/\zeta_d)^{\frac{1}{2}}}. \tag{22}$$

The temperature dependence of λ_u/λ_d will mainly depend on W′, the energy difference between the n^- levels and the upper band. Roughly

$$\lambda_u/\lambda_d \propto e^{-\mathrm{W}'/k\mathrm{T}}. \tag{23}$$

The temperature variation of $(\sigma_{\text{light}})_{\text{sat}}$ itself is given by

$$(\sigma_{\text{light}})_{\text{sat}} \propto e^{-\mathrm{W}'/k\mathrm{T}}. \tag{24}$$

If W′ is fairly small, as it well may be, this equation shows that $(\sigma_{\text{light}})_{\text{sat}}$ may only vary in a trivial way with the temperature unlike the large variations of σ_{dark}.

More important than the light saturated conductivity is the extra conductivity for small light intensity. It is easily verified that this is given by the formula

$$\delta\sigma \propto \delta u \sim \tfrac{1}{2} \frac{\kappa \mathrm{I}}{\lambda_d} \frac{1}{(\zeta_u/\zeta_d)^{\frac{1}{2}}}, \tag{25}$$

or

$$\delta\sigma \propto \frac{\mathrm{I}}{\sigma_{\text{dark}}} \tag{26}$$

approximately. This is a very striking temperature variation, but it holds only so long as the extra conductivity is small compared with the dark conductivity. For larger values of I such that the conductivity is large compared with the dark conductivity but at the same time small compared with the saturated light conductivity, we can neglect all the ζ-terms and also n compared with λ_u/λ_d in (20), obtaining

$$\rho_d n^3 + \lambda_u n^2 = \kappa \mathrm{I} \lambda_u/\lambda_d. \tag{27}$$

This gives

$$\sigma_{\text{light}} \propto n \propto \mathrm{I}^{\frac{1}{3}}, \tag{28}$$

and σ_{light} is roughly independent of the temperature in this range.

The variety of behaviour found for the proposed model may prove to represent the facts for some substances, but, if the facts for steady illumination as at present recorded can be relied upon to apply to the electrodeless conduction here discussed, it seems certain that in some substances at least the extra conductivity is proportional to the illumination over a wide range and does not satisfy (28) in the range to which the theory should apply. For such substances the model must be modified.

§ 8. *A Modified Model.*—Let us assume that a primary function of the light is to excite electrons from the " negative " polar impurities to the upper band of running levels. The same light may or may not play a part in continually creating the polar impurities. The only modification of the model is that we must replace λ_u by $\lambda_u + \kappa'\mathrm{I}$ and possibly may then neglect λ_u. Equation (20) becomes, neglecting λ_u,

$$\frac{\rho_d\kappa'}{\lambda_d}\,\mathrm{I}\left(n^2 - \frac{\zeta_u}{\zeta_d}\right)^2 + \kappa'\mathrm{I}n\left(\frac{\kappa'\mathrm{I}}{\lambda_d} - n\right)\left(n^2 - \frac{\zeta_u}{\zeta_d}\right) = (\kappa\mathrm{I} + \rho)\,n\left(\frac{\kappa'\mathrm{I}}{\lambda_d} - n\right)^2. \quad (29)$$

Under certain conditions when I is fairly large, we may repeat again the solution (21), which gives $\sigma \propto n \propto \mathrm{I}$. These conditions are rather complicated and require that $\kappa \gg \kappa'$ so that the right-hand term in (29) remains dominant unless $n \sim \kappa'\mathrm{I}/\lambda_d$. We shall not investigate further the large variety of situations here possible. It is sufficient to have shown that $\sigma \propto \mathrm{I}$ is not impossible.†

Quite other types of model must certainly be used also. The rock salt type of photoconductivity as ordinarily investigated for the yellow form depends on the presence of atomic impurities which are not being steadily created during the experiment. To deal with this case, one can obviously set up other modified models in which allowance may or may not be made for a slow destruction of the atomic impurities during the illumination. But until the facts corresponding to the theory have been determined further elaboration is out of place.

§ 9. *Rate of Approach to Equilibrium.*—There remains one point which can be investigated in a very general way for all models of the type here contemplated, so that perhaps its consideration is justified. If a steady state has been attained for a given value of I and I is then changed to a new constant value, how quickly will the new equilibrium state be approached ? We shall be able to show that the rate of approach to a large conductivity (σ_{light}) is much more rapid than the approach to the small σ_{dark} ; this is a general feature of our models and appears to be in good agreement with the facts for selenium and zincblende.†

---

† *Cf.* Teichmann, 'Proc. Roy. Soc.,' A, vol. 139, p. 105 (1933).

If we write $n = \bar{n} + \delta$, etc., in equations (11) to (14) and study the way in which the δ's tend to zero, it is easily seen that we shall obtain three equations of the form

$$\frac{d\delta}{dt} = A_1\delta + B_1\delta^- + C_1\delta^+, \tag{30}$$

$$\frac{d\delta^-}{dt} = A_2\delta + B_2\delta^- + C_2\delta^+, \tag{31}$$

$$\frac{d\delta^+}{dt} = A_3\delta + B_3\delta^- + C_3\delta^+, \tag{32}$$

in which in general the coefficients A_1, ..., C_3 are linear functions of the equilibrium concentrations $\bar{n}$. The solutions of these equations must, of course, be of the well-known form $\Sigma_1^3\,\alpha_s e^{-\lambda_s t}$ where the λ's are the roots of the cubic determinantal equation. Now since all the A_1, .., C_3 are linear functions of the $\bar{n}$'s, they will in general for large increases of the $\bar{n}$'s by illumination be roughly proportional to the $\bar{n}$'s, and so also in general will the λ's. *Thus the rate of approach to any equilibrium state will be proportional* (*roughly*) *to the concentrations of free electrons, etc., in that state.* It is an observed fact that this rate is much faster in the light than in the dark, for example, for selenium, and it is possible that the model has here suggested the correct explanation.

Summary.

The paper is a critical survey of certain current views on the nature of crystalline insulators and their electronic properties. In particular Tamm's theory of surface levels on an insulator is analysed and is shown not to be of importance in the theory of the passage of current through the insulator. It is shown that the necessary charges at an interface between a metal and a crystal can always be accommodated in the metal at levels very close to the top of the Fermi-Sommerfeld distribution. The manner in which a poorly conducting crystal carries a current between metal electrodes is also discussed, and the results of the discussion applied to the properties of crystals rendered conducting by light. It is suggested that certain experiments on photoconductivity should be carried out for constant illumination by a technique which uses no electrodes and creates no space charges in the crystal. The type of theory which should apply to such cases is briefly outlined.

---

† Nix, 'Rev. Mod. Physics,' vol. 4, p. 723 (1932) esp. p. 739. See also Hughes and Du Bridge, *loc. cit.*, p. 323 (molybdenite).

The Relationship between Viscosity, Elasticity and Plastic Strength of a Soft Material as Illustrated by some Mechanical Properties of Flour Dough.—III.

By ROBERT KENWORTHY SCHOFIELD and GEORGE WILLIAM SCOTT BLAIR.

(Communicated by Sir John Russell, F.R.S.—Received January 2, 1933.)

The equation

$$\eta = n \cdot t_r \qquad (1)$$

connects the viscosity, η, with the rigidity modulus, n, and the time of relaxation, t_r. In the case of simple fluids, for which the equation was originally given by Maxwell, the dissipation of shearing stress is so rapid that neither n nor t_r has been measured, so that the relationship has only a theoretical interest.

In the first paper of this series* it was shown that in flour dough stress dissipation proceeds at rates that can readily be followed experimentally The " time of relaxation," t_r, was given an extended significance, and defined in such a way as to be applicable to plastic materials like dough for which η is not a constant. A series of values for the viscosity was obtained with the aid of equation (1).

In the second paper† a description was given of a method by which a record or " rheogram " can be obtained showing the amount of flow that has occurred in a given time under stress. In this way another series of viscosity values was obtained. A comparison of the two sets showed a satisfactory agreement so far as order of magnitude was concerned, but it revealed quantitative discrepancies which made a further investigation desirable.

In both papers the fact that the viscosity is dependent on the total strain as well as on the stress was the subject of comment. Flour dough thus exhibits a behaviour reminiscent of " work-hardening " in metals, and it is evident that we can only expect agreement between values of η in cases where the " history " of the specimens has been the same. Further study has revealed the existence of two more effects which have close parallels in the behaviour of metals, namely, elastic after-effect and elastic hysteresis.‡ It is shown below that by taking due account of these properties the seemingly conflicting results of Paper II fall into line.

---

* 'Proc. Roy. Soc.,' A, vol. 138, p. 707 (1932), referred to as Paper I.

† 'Proc. Roy. Soc.,' A, vol. 139, p. 557 (1933), referred to as Paper II.

‡ *Cf.* Nadai, " Plasticity," McGraw Hill Book Co., New York (1932).

Elastic After-effect.

The apparatus was a modification of that used to follow the relaxation of stress and described in Paper I. A cylinder of dough, about 15 cm. long and 0·7 cm. diameter, was floated on a mercury bath. Two small scales, graduated in quarter centimetres and tenths, were attached one at either end of the dough by means of cork " chairs " which adhere readily to it To one scale (scale A) was fastened a thin strand of rubber, about 20 cm long, the other end being anchored to a stout pin sticking up in the trough. To the other scale (scale B) was attached a piece of sewing cotton, the other end of which was wound on a small winch controlled by a worm and crank. Two low-power microscopes were trained one on either scale. One-tenth part of each small division could be estimated, so that a movement of 0·0025 cm. could be detected with certainty, fig. 1.

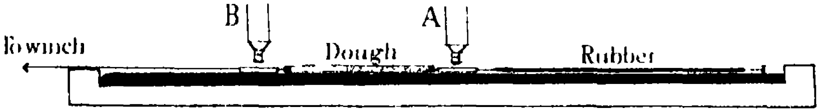


FIG. 1.

The position of scale A is a direct indication of the stress. Calibration was effected by placing the trough erect, when empty, and hanging weights on the lower end of the scale, due allowance being made for the weight of the scale and rubber in obtaining the zero. The elastic properties of the rubber, although not perfect, were found to be sufficiently satisfactory. For extensions up to 10 per cent. a constant factor of 690 dynes per scale division was found. To obtain the extension of the dough cylinder, the shift of scale A must be subtracted from the corresponding shift of scale B.

The doughs were made up in the way described in Paper II, the moisture content being so adjusted that the dough did not stick appreciably to a glass plate when pressed firmly against it. The doughs were allowed to stand for 1½ hours before use, and the cylinders were formed by extrusion from a " gun." The part of the mercury trough occupied by the cylinder was flanked by a strip of wet felt and covered by a glass-topped frame to minimize drying.

When the tensile stress on a dough cylinder is increased, a point is reached at which the plastic strength is exceeded and plastic (non-recoverable) extension occurs. If the tensile stress is below the plastic limit, an extension, due to a small increase in stress, appears to be wholly recoverable on restoring the stress to its former value, provided that sufficient time is allowed. Fig. 2, *a* and *b*, is the record of such an experiment. The dough had been under a stress of 4700 dynes/cm.^{2} for some minutes (the plastic limit in this case being about 5000 dynes/cm.^{2}) and neither scale was moving appreciably. The handle of the winch was then given a turn which caused the scale B to move 0·41 divisions.

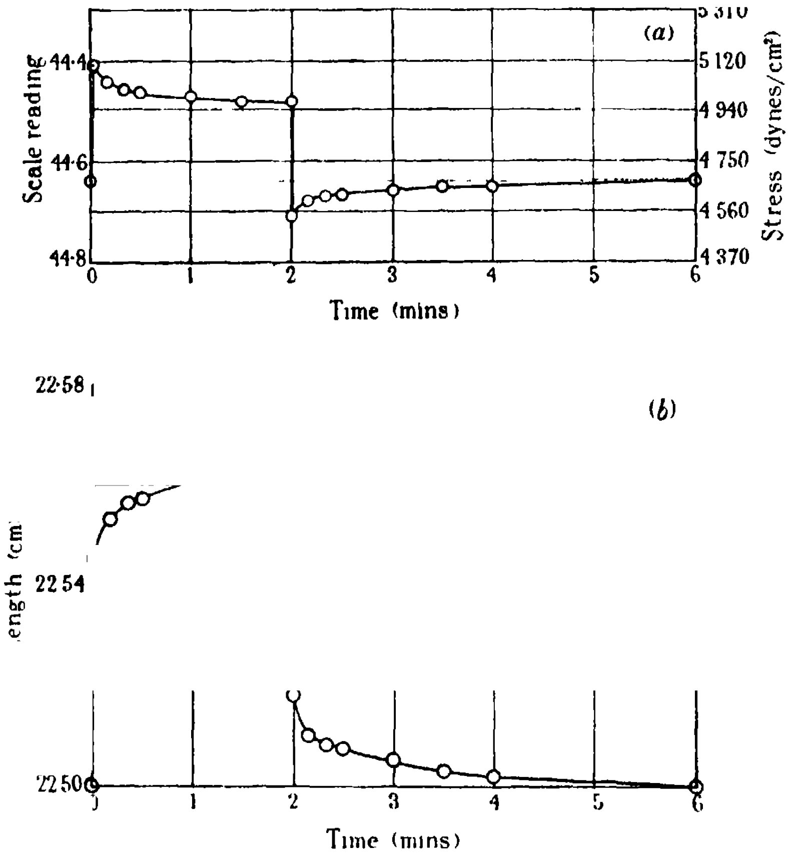


FIG. 2.

After two minutes a reverse turn was given to the crank which restored the scale to its original position. Fig. 2, *a* shows the movement of scale A which, as already noted, is a direct indicator of the stress, and of which the shift must be subtracted from 0·41 (the shift of scale B) during the first 2 minutes, and from zero subsequently, in obtaining the elongations shown in fig. 2, *b*. It will be seen that the value obtained for the rigidity modulus by dividing one-third of the change in tensile stress by the corresponding elongation, depends materially on the time allowed. If the first reading obtained after turning the winch is used, the value is $n = 7 \cdot 7 \times 10^4$; after 2 minutes it is $n = 4 \cdot 3 \times 10^4$. It will be seen that both the stress and the strain returned to their original value so that no perceptible plastic flow occured during the period occupied by the experiment. The phenomenon exhibited is essentially that of elastic after-effect.

Fig. 3 shows the same thing at a higher stress, namely, 5400 dynes/cm.^{2}. Here the effects of slow plastic flow and of elastic after-effect are superimposed.

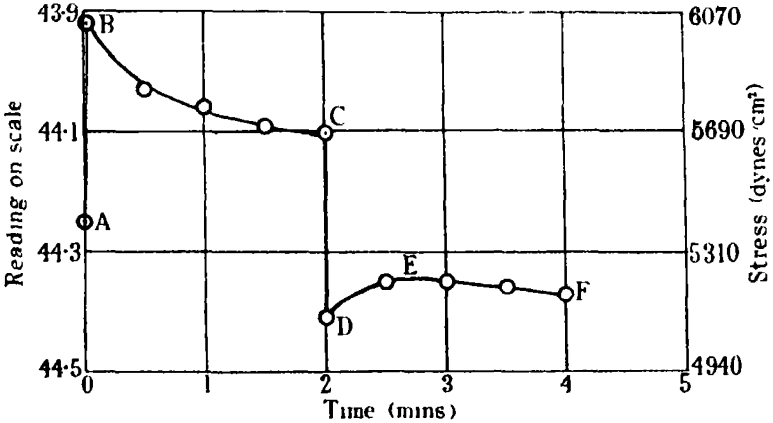


FIG. 3.

Elastic after-effect was also studied by first extending a dough cylinder at a uniform rate by driving the winch from a suitably geared electric motor and then suddenly releasing the stress altogether by burning the cotton. Fig. 4 shows the variation of length with time in a typical experiment. According to the Maxwellian theory, the rate of fractional elongation de/dt is related to the shearing stress, S (which is one-third the tensile stress, s, for this material), thus

$$\frac{de}{dt} = \frac{1}{n}\frac{dS}{dt} + \frac{1}{\eta} \, . \, S. \qquad \text{(ii)}$$

At the point B in fig. 4 the stress was suddenly reduced to zero. Equation (ii) indicates that a large negative de/dt is to be expected momentarily, but that de/dt should immediately afterwards become zero. Actually de/dt maintained an appreciable negative value for many minutes during which both S and

dS/dt were zero. Such a case can only be covered by an extension of equation (ii), such as

$$\frac{de}{dt} = \left(\frac{1}{n} \cdot \frac{dS}{dt} - \frac{d\alpha}{dt}\right) + \frac{1}{\eta} \cdot S, \tag{iii}$$

in which α represents the influence of elastic after-effect.

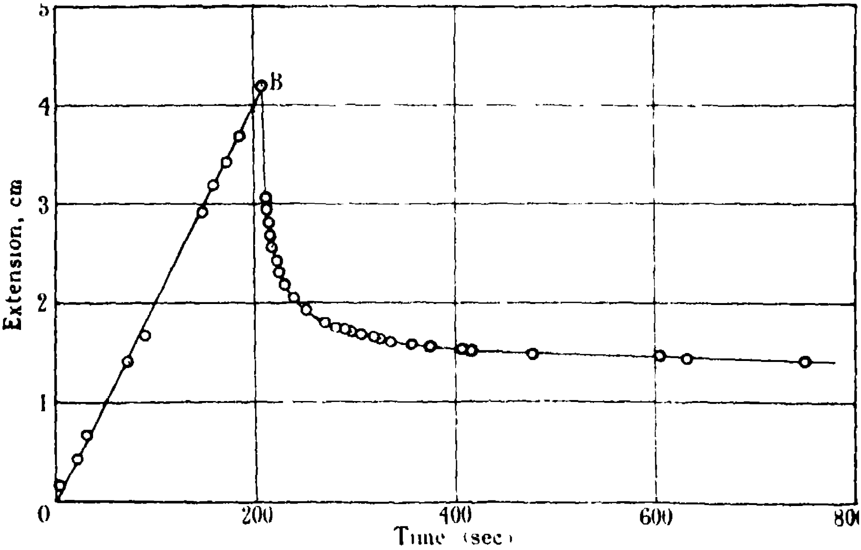


FIG. 4.

Applying this relation to the case illustrated in fig 2, $\frac{1}{\eta} \cdot S$, which represen the rate of plastic flow, may be neglected, and we obtain by integration

$$\Delta e = \frac{1}{n} \Delta S - \Delta\alpha.$$

As a steady condition is approached when the system is left undisturbed, it is reasonable to regard α as tending to zero under these circumstances. Interpreted in this light, fig. 2 indicates that on rapidly increasing the stress, α assumes a positive value which falls gradually towards zero so long as there is no further abrupt change in stress. Conversely, a rapid decrease in stress imparts a temporary negative value to α.

The fact already noted that in fig. 3 the effects of plastic flow appear simply to be superimposed on the elastic effects shown in fig. 2, is reflected in the form of equation (iii) which shows de/dt to be the simple sum of an elastic part (within the bracket) and a plastic part. A particularly interesting feature of fig. 3 is the spontaneous rise of the stress to a maximum from D to E. An unthinking use of the expression

$$t_r = -\,\mathrm{S}\Big/\frac{d\mathrm{S}}{dt} \text{ (at constant elongation)} \tag{iv}$$

would suggest that the relaxation time had a negative value in this region. Reference to equation (iii) shows, however, that when de/dt is zero, $t_r = \eta/n$ is only equal to $-\,\mathrm{S}\Big/\frac{d\mathrm{S}}{dt}$ when $d\alpha/dt$ is zero; and this only occurs when no rapid change of stress has taken place for some time. The definition of t_r, adopted in Paper I and embodied in equation (iv), is therefore subject to this limitation. Furthermore, of the two values for n, given in connection with fig. 2 for the modulus, it is now clear that the lower one, $4 \cdot 3 \times 10^4$, must be preferred when relating η to t_r.

Elastic Hysteresis.

An experiment such as that of fig. 4 can also give a value for n if, in addition to a determination of the total contraction, a measurement is made of the stress just before burning the cotton. Table I summarizes the results of three such experiments. The striking fact to be noted is the smallness of the values found for n in comparison with that deduced above from fig 2. The difference cannot be put down to elastic after-effect.

Table I.

s.	l_1.	l_2.	n.
5200	18·1	16 0	1 40 × 10^4
5400	19·7	17·4	1·45 × 10^4
5200	20 3	18·1	1·51 × 10^4

s is the tensile stress, n, the rigidity modulus, and l_1 and l_2 the lengths of the dough under stress, and after recovery respectively.

The origin of this discrepancy became apparent from a further experiment. A dough cylinder was first extended at a slow uniform rate by about 2·5 cm. so as to remove any tendency to curl, and was then held extended for about 5 minutes to allow the stress to fall to a value in the neighbourhood of the plastic limit, at the end of which time its rate of decrease was quite slow. This condition is indicated by the point A on fig. 5.

The stress was then released by stages. Fifteen seconds was allowed between each turn of the crank, the stress recorded being that at the end of each 15-seconds period. It was necessary to fix some time interval of this kind in order that the small influence of elastic after-effect should be spread evenly over the experiment. At B nearly all the stress had been released and it was then increased in the same way to C; and the cycle twice repeated. The lateral shift of the loop at each repeat is evidence that plastic flow occurred at the higher stresses. It appears safe to assume, however, that this complication does not arise in the lower portion of each loop. In following any *continuous* curve it is found that the modulus connecting the change of stress with the corresponding change of length always *decreases*, but that at each discontinuity where the sense of the stress change alters, the modulus abruptly *increases* very considerably. At the points B, D and F the modulus changes from about 1 × 10^4 to 4 × 10^4.

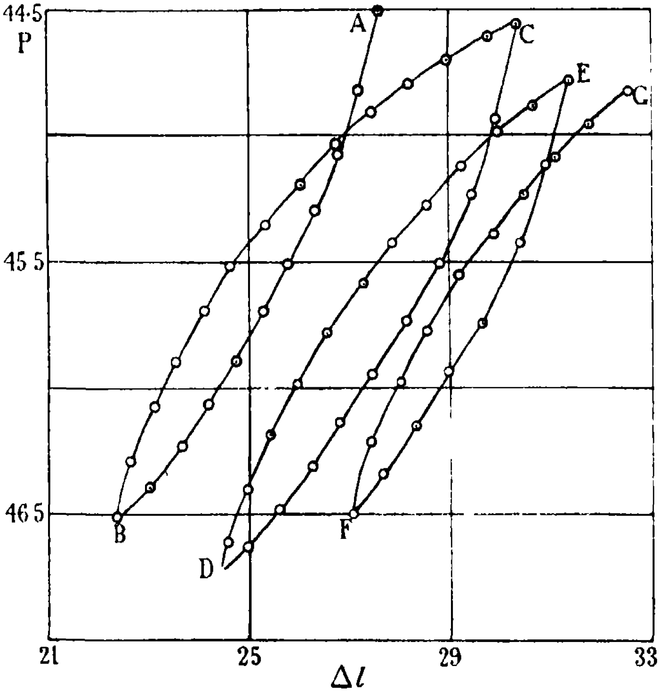


FIG. 5.

The modulus obtained from fig. 2, using 15 seconds as the time interval, should evidently be compared with the highest values obtained from fig. 4 since both refer to the condition just after a change in sense of stress variation.

The moduli quoted in Table I, as well as those cited in the earlier papers of this series, are in a sense mean values corresponding to a change such as that from A to B, C to D, or E to F on fig. 5, and we should expect these values to be small. The exact magnitude of such a mean value will evidently depend, among other things, on the magnitude of the stress change, and will tend to decrease as this increases. The recognition of this phenomenon, which is essentially a case of elastic hysteresis, enables a rational explanation to be formulated for the marked inconstancy of the modulus It also makes it necessary to reconsider the value previously adopted for the modulus when deriving the viscosity from the relaxation time.

Test of the relation $\eta = nt_r$.

The presence of elastic after-effect and elastic hysteresis adds considerably to the difficulty of devising an experimental check on equation (1). The only satisfactory way would appear to be to measure all three quantities in rapid succession on the same piece of dough. A number of experiments of this kind were carried out with the apparatus slightly rearranged in the manner shown in fig. 6. A scale (scale C) was inserted at the right-hand end of the rubber strand and held by a cotton wound on a hand-controlled winch. The winch at the left-hand end (off the diagram) was connected to an electric motor through a suitable gearing provided with a "clutch" by means of which the winding could be started or stopped at will.

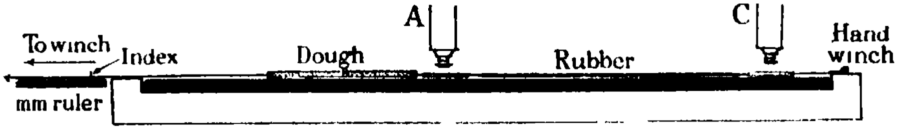


FIG. 6

During the first part of each experiment the motor-driven winch was allowed to pull the dough out slowly, the rate of extension being determined by observing the motion of an index on the cotton as it moved over a suitably placed millimetre ruler. At the same time, scale A was kept under observation. With this arrangement the stress is given by the difference of the readings of scales A and C, but since during this part of the experiment the position of scale C was not altered, the movement of scale A was a direct indication of the rise in the stress. After a while the increase of stress became very slow, and then when a suitable division of scale A exactly coincided with the cross-wire of the observing microscope, the clutch was thrown out, the time being noted. Simultaneously the hand-winch was released by one turn. This caused both the scales to shift. Scale C moved in direct response a distance of 0·19 scale divisions; scale A first moved rapidly to the left and then more slowly returned towards its former position. As soon as it reached its old position the crank was given a further turn, the time being again noted. The process was repeated a number of times. The method used here is essentially the same as that adopted earlier in following the relaxation of stress and already described in Paper I, but there is one important addition. With the former apparatus the movement of the end of the dough attached to the rubber strand could be observed but not measured. By *measuring* the extent of the rapid *elastic* reaction of the dough to the movement of the hand-winch a value can be obtained for the modulus The evaluation of η, t_r and n is most readily explained by taking a single experiment and working it out in detail.

Evaluation of η.—In a typical case the dough was first pulled out at a steady rate of 0·0154 cm. per second until its length was 25·3 cm. At this point $\frac{de}{dt} = \frac{0 \cdot 0154}{25 \cdot 3} = 6 \cdot 1 \times 10^{-4}$ sec.$^{-1}$. Although the pull of the rubber strand had become steady, the cross-section of the dough was diminishing, so that the stress per unit area was rising. From the geometry of the case, $\frac{1}{S}\frac{dS}{dt} = \frac{de}{dt}$; hence the elastic part of the elongation, namely, $\frac{1}{n}\frac{dS}{dt}$, equals $\frac{de}{dt} \times \frac{S}{n}$. S, the shearing stress which is one-third the tensile stress per unit area of cross-section, was found from the scale readings and the dimensions of the cylinder to be $2 \cdot 7 \times 10^{3}$ dynes/cm.^{2}. The appropriate value for n we do not know exactly, but as the stress had been slowly rising from zero it was certainly low and probably not greater than 1×10^{4}. This gives $\frac{1}{n}\frac{dS}{dt} = 1 \cdot 6 \times 10^{-4}$ sec.$^{-1}$. Owing to the comparatively steady stress conditions, $d\alpha/dt$ of equation (iii) will be negligible; hence the rate of plastic flow is simply

$$(6 \cdot 1 - 1 \cdot 6) \times 10^{-4} = 4 \cdot 5 \times 10^{-4} \text{ sec.}^{-1}.$$

Dividing into S we have

$$\eta = \frac{2 \cdot 7 \times 10^{3}}{4 \cdot 5 \times 10^{-4}} = 0 \cdot 6 \times 10^{7} \text{ dynes . cm.}^{-2} \text{. secs.}$$

Evaluation of t_r.—Equation (iv) for t_r holds as we have seen when there is no appreciable change in α. As soon as the clutch is thrown out the stress starts to fall, somewhat rapidly at first and more slowly afterwards. After the first turn of the hand-winch 9 seconds elapsed before scale A returned to its original position. The stress change as given by the shift of scale C was 0·19 scale divisions and the mean stress in the same units was 2·67, hence

$$S\Big/\frac{dS}{dt} = \frac{2\cdot67 \times 9}{0\cdot19} = 126 \text{ secs.}$$

As α must have been taking on an increasing negative value during this time, this figure will be a little below the true value for t_r. Reference to fig. 2 indicates that 0·19 scale divisions in 9 seconds is not an excessive rate, and suggests that the discrepancy will not be large. The series of values thus obtained were

126, 275, 445, 760, 1280, 1980 seconds.

These are mean values, and an estimated value of

$$t_r = 100 \text{ seconds}$$

for the relaxation time immediately after throwing out the clutch seems reasonable.

Evaluation of n.—In calculating n from the rapid elastic reaction of scale A to the movement of the hand-winch, the influence of elastic after-effect is more serious. After the turn of the crank which immediately followed the throwing out of the clutch, the leftward movement of the scale was arrested by plastic flow in the dough after about 1 second, a time interval which must have been insufficient to allow the full elastic movement to occur. After the sixth turn more than 20 seconds elapsed before the scale resumed its rightward movement. It is not therefore surprising to find that n apparently changes from 11×10^4 at the first shift to 6×10^4 at the sixth. Owing to elastic hysteresis some fall in n is to be expected, but there can be little doubt that the second figure is nearer the true value of n immediately after the clutch was thrown out. This conclusion is confirmed by the value

$$n = 6 \times 10^4 \text{ dynes/cm.}^2$$

obtained from the ratio of the estimated values of η and t_r. The closeness of the agreement is partly fortuitous as round figures have been used throughout. In some of the other experiments of the same kind the concordance was not so close, but no case was found which, when due allowance had been made for the numerous uncertainties, provided clear evidence in conflict with the equation

$$\eta = n \,.\, t_r. \qquad \text{(i)}$$

In the last analysis it appears doubtful whether this equation is capable of direct proof, seeing that η, n and t_r are all variable quantities, and it does not appear possible to measure all three actually simultaneously. It is perhaps true to regard equation (i) as defining t_r in terms of η and n which in their turn are defined by equation (iii). We may then take the above experiments as indicating that by using equation (iii) we do not obtain values for η and n which show inconsistant or unreasonable fluctuations.

The lack of agreement between the values obtained for η from the rheograms in Paper II and those calculated from n and t_r using the data of Paper I is no longer surprising Undoubtedly the chief cause of the discrepancy is to be found in the value 2×10^4 used throughout for n. This is certainly a good *mean* value for the doughs used, but it is now clear that the conditions under which t_r is measured correspond to a part of the hysteresis loop in which n has a value appreciably above the mean. A value in the neighbourhood of 5×10^4 would accord better with the findings of this paper and also bring the two sets of figures into line. It will be noted that the viscosity found above, namely, $0{\cdot}6 \times 10^7$, is lower than the general run of the values given in the table in Paper II. This is to be expected since the tensile stress was 8100 dynes/cm.^{2}, which is higher than for any viscosity included in the table; the extension in the units used was $0{\cdot}43$.

Conclusions.

Taken together, our experiments seem to show that the mechanical properties of flour dough can be described by the equation

$$\frac{de}{dt} = \left(\frac{1}{n}\frac{dS}{dt} - \frac{d\alpha}{dt}\right) + \frac{1}{\eta} . S. \tag{iii}$$

n varies in a manner characteristic of elastic hysteresis. Consequently it is liable to increase abruptly when dS/dt changes sign, and is generally dependent on the history of the specimen. Direct measurements of n have been only possible at stresses below the limit of plastic strength, but indirect determinations obtained by dividing η by t_r have been made above it, and there is no evidence of any change in n in passing across this limit other than that to be anticipated from hysteresis.

α, as we have seen, is only of importance when abrupt changes have recently occurred in S. Like n, it does not appear to depend on the absolute value of S, but rather on its value relative to those which have obtained in the immediate past.

η shows quite a different behaviour. It is extremely sensitive to the absolute value of S. So rapidly does it increase as S is reduced, that only a narrow margin of stress separates the condition in which plastic flow is dominant from that in which it can scarcely be detected. Herein lies the justification for the use of the words "plastic limit" and "plastic strength." For, although a *precise* limit could only be defined by arbitrarily specifying a certain viscosity as marking the limit, no serious misunderstanding is likely to arise in practice. Furthermore, although η is strongly influenced by the previous history, it appears to be affected by factors quite distinct from those which control n. As we have seen, n is largely determined by the previous movements of S; η, on the other hand, is mainly affected by the amount of flow that has taken place, and increases for a given value of S as the elongation proceeds. A series of experiments, of which the results are shown in fig 7, showed clearly

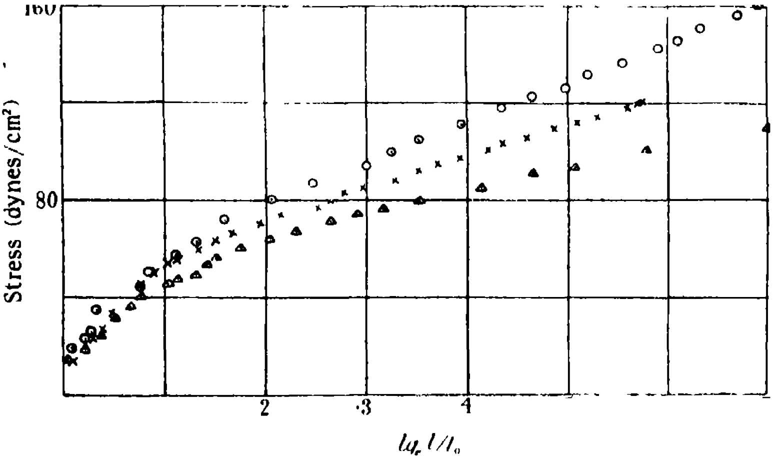


Fig. 7.

that the extent of this hardening is controlled by the amount of the elongation as well as by its rate. The apparatus as shown in fig. 1 was used, and the winch was driven at a steady rate by the electric motor, a different speed being used in each experiment. The similarity of the curves, for which the extreme speeds were as two to five, provides clear evidence on this point. It is interesting to note how similar is the behaviour exhibited in fig. 7 to that shown by soft metals, particularly at high temperature; the same can equally be said of fig. 5 (Nadai, *loc. cit.*).

In conclusion, an electrical analogy may prove helpful to some when considering the meaning of equation (iii), and its use in the various special cases considered above. Using the formal analogy which exists between viscosity and electrical resistance, and between elasticity and electrical capacity, it may be said that the electrical behaviour of the system sketched in fig. 8 reproduces the essential mechanical properties of dough. Fig. 8, *a* represents the condition after the point B of fig. 4 at which the stress on the dough is suddenly released. In the electrical case it is supposed that an e.m.f. was first applied between A and B and that the system was then short-circuited. The current in R would cease at once and the condenser, C, would instantly discharge. The charge in *c* would take some time to leak through the high resistance, *r*. The branch *cr* is evidently responsible for the "after-effect" and will give a finite "α" whenever the e.m.f. across C differs from that across *c*, which will be the case after any abrupt change in the external e.m.f. In fig. 8, *b* the system after being subjected to an e.m.f. has been disconnected. This condition corresponds with stress relaxation at constant elongation. The condensers discharge through R; unless R is large compared with *r*, "α" will make its appearance. In fig. 8, *c* we have the analogy to conditions which furnished data for fig. 7, namely, the steady application of stress. To make the analogies closer, the resistances and capacities should be considered to be variable.

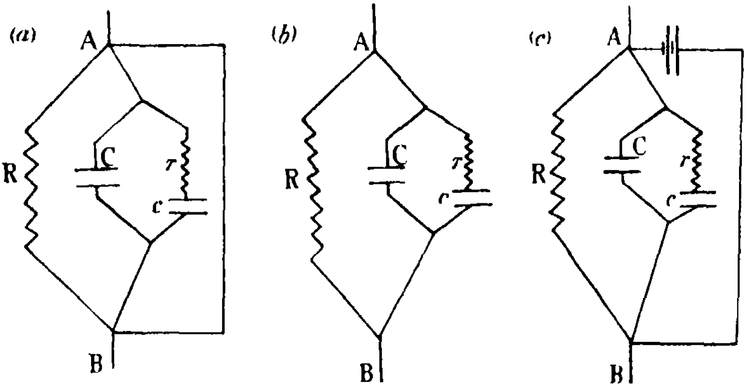


FIG. 8.

There is, however, an important contrast, in that the three elements which go to make up the mechanical system must be regarded as linked end to end and not "in parallel" as in the electrical case. The elastic elements in the dough, the behaviour of which is described by the symbols *n* and α, are presumably formed by the protein constituents, while the plasticity embodied in the symbol

η must be associated with the linkage between the elastic elements. The individual elastic elements appear to possess mechanical properties similar to those which Shorter and others* have found in hairs.

Summary.

A further study of the mechanical properties of flour dough has revealed the presence of two properties in addition to hardening, both of which are well known in the study of metals ; namely, elastic after-effect and elastic hysteresis.

The first necessitates the addition of a term $d\alpha/dt$ to the Maxwell equation. which then becomes

$$\frac{de}{dt} = \left(\frac{1}{n}\frac{dS}{dt} - \frac{d\alpha}{dt}\right) + \frac{1}{\eta}\ S.$$

This term is only important when abrupt changes of stress have recently occurred.

The second property causes n to decrease steadily whenever dS/dt preserves the same sign for some time, and to increase abruptly when the sign of dS/dt is changed.

In Paper II it was shown that the viscosity, as determined from the rate of flow, agreed roughly, but not exactly, with that calculated as the product of the rigidity modulus and the relaxation time. It is now clear that the value adopted for n was a mean value, and differed somewhat from that appropriate to the conditions during stress relaxation. Due appreciation of this point renders the agreement quantitative.

---

* 'J. Text. Inst.,' vol. 15, p. 207 (1924); 'Trans. Faraday Soc.,' vol. 20, p. 228 (1924); 'J. Soc. Dyer Col. (Bradford),' vol. 41, p. 212 (1925).

The Behaviour of Electrolytes in Mixed Solvents. Part V.—The Free Energy of Lithium Chloride in Water-Alcohol Mixtures and the Salting-out of Alcohol.

By J. A. V. BUTLER, D.Sc., and D. W. THOMSON, Ph.D., University of Edinburgh.

(Communicated by J Kendall, F R.S —Received January 25, 1933.)

The Salting-out of Alcohol by Lithium Chloride.

In Part II of this series of papers* are recorded measurements of the partial vapour pressures of solutions of lithium chloride in water-ethyl alcohol mixtures, the alcohol contents of which extended from 6·4 to 100 mols. per cent. It has seemed desirable to extend this range so as to include some smaller concentrations of alcohol, and accordingly measurements have now been made of solutions containing 2 and 4 mols. per cent of alcohol and in each case 0·5, 1·0 and 4·0 m. lithium chloride. The experimental method was the same as that previously described, except that in these cases the viscosity method was less suitable for determining the composition of the condensate, which was obtained by comparison with known compositions in the interferometer The final values of the partial pressures, each being the mean of at least two determinations, are given in Table I. α_w/α_w^0 and α_a/α_a^0 are the relative activities, *i.e.*, the ratio of the partial pressure in a given solution to the partial pressure

Table I.—Partial vapour pressures and activities of water and alcohol in water-alcohol solutions of lithium chloride.

N_a.	m.	p_w	p_a.	a_w/a_w^0.	a_a/a_a^0.
2·0	0·0	23·00	4·08	1·000	1·000
	0·5	22·73	4·48	0·988	1·098
	1·0	22·14	4·83	0·963	1·183
	4·0	18·40	6·45	0·800	1·581
4·0	0·0	22·40	8·16	1·000	1·000
	0·5	22·11	8·84	0·987	1·083
	1·0	21·58	9·36	0·964	1·148
	2·0	20·52	10·42	0·915	1·276
	4·0	17·71	12·11	0·791	1·485
6·4	0·0	21·91	12·32	1·000	1·000
	4·0	16·64	18·03	0·760	1·462

---

* ‘Proc. Roy. Soc.,’ A, vol. 129, p. 519 (1930).

of the corresponding solvent. Two of the solutions containing 6·4 mols. per cent. alcohol were also redetermined and gave values in close agreement with those previously obtained.

We are now in a position to discuss in more detail the behaviour of these solutions. It was shown in the previous paper that while the salt increased the partial pressure of alcohol in solutions containing smaller proportions of alcohol, it caused a lowering of the pressure in the more alcoholic solutions, but the relative lowering of the vapour pressure of alcohol was always less than that of water, indicating that in all solutions the alcohol was salted out with respect to the water In order to assess this effect more precisely it is necessary to consider what will be the behaviour of a solute which interacts equally with the two solvent molecules.

We will consider first a solution of a solvent S_1 and a binary salt S_3. If n_1, n_3 are their amounts in mols. and $\bar{F}_1$, $\bar{F}_3$, their partial molar free energies, the variations of the latter at constant temperature and pressure are conditioned by

$$n_1 d\bar{F}_1 + n_3 d\bar{F}_3 = 0. \tag{1}$$

When n_3 is small compared with n_1, the variation of $\bar{F}_3$ is given by

$$\bar{F}_3 = \bar{F}_3^\circ + 2\mathrm{RT} \log n_3,$$

so that

$$d\bar{F}_1/dn_3 = -2\mathrm{RT}/n_1. \tag{2}$$

In the case of a mixed solvent containing the components S_1, S_2 and the binary salt S_3, we have

$$n_1 d\bar{F}_1 + n_2 d\bar{F}_2 + n_3 d\bar{F}_3 = 0, \tag{3}$$

and if n_3 is small the variation of $\bar{F}_3$ is again given by $\bar{F}_3 = \bar{F}_3^0 + 2\mathrm{RT} \log n_3$, so that

$$n_1 (d\bar{F}_1/dn_3) + n_2 (d\bar{F}_2/dn_3) = -2\mathrm{RT}, \tag{4}$$

or dividing by $(n_1 + n_2)$,

$$\mathrm{N}_1 (d\bar{F}_1/dn_3) + \mathrm{N}_2 (d\bar{F}_2/dn_3) = -\frac{2\mathrm{RT}}{n_1 + n_2}, \tag{5}$$

where $\mathrm{N}_1 = n_1/n_1 + n_2$, $\mathrm{N}_2 = n_2/n_1 + n_2$ are the molar fractions of the components of the solvent. If the solute S_3 has an equal effect on both of the components of the solvent, we must suppose that

$$\frac{d\bar{F}_1}{dn_3} = \frac{d\bar{F}_2}{dn_3} = -\frac{2\mathrm{RT}}{n_1 + n_2}, \tag{6}$$

or writing,

$$d\bar{F}_1 = RT\, d\log(p_1/p_1^0), \quad d\bar{F}_2 = RT\, d\log(p_2/p_2^0),$$

the vapour pressure lowerings of the components of a binary solvent produced by a salt which interacts equally with both components are given by the expressions

$$\frac{d\log(p_1/p_1^0)}{dn_3} = \frac{d\log(p_2/p_2^0)}{dn_3} = -\frac{2}{n_1+n_2} \qquad (7)$$

We shall regard this as the " normal " behaviour of a salt in a mixed solvent, neither component of which is salted out.

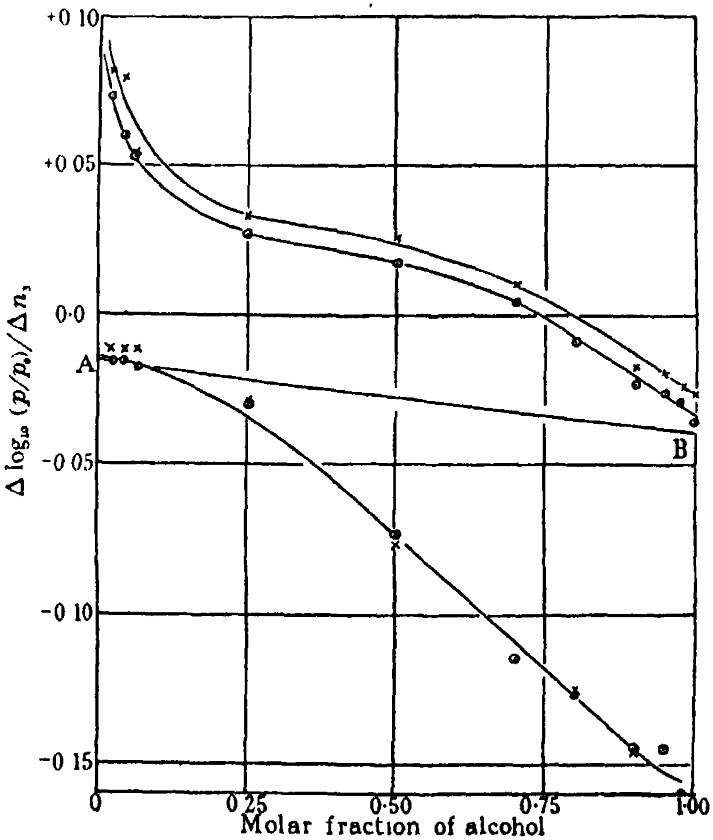


FIG. 1.—Effect of lithium chloride on the partial pressures of alcohol and water. (Upper curves, alcohol; lower curve, water) —⊙— 1 m. LiCl. —×— 0·5 m. LiCl.

In fig. 1 the line AB shows the molecular lowerings of the vapour pressure in the normal case, expressed as $d\log_{10}(p/p^0)/dn_3$ and calculated by (7) for a binary salt in water-alcohol mixtures. The experimental values of $\Delta \log_{10} (p/p^0)/\Delta n_3$ evaluated for the intervals 0–0·5 m. and 0–1·0 m. are also shown. These values are considerably displaced from the normal curve, the alcohol positively and the water negatively. Equation (7) is only strictly applicable to the normal case at very small salt concentrations, and a correction for the activity coefficient of the salt would be necessary for a large concentration interval. But the observed displacements are much greater than can be accounted for in this way, and it is necessary to conclude that alcohol is salted out from the whole range of solutions.

The extent of the salting out is properly measured, not by the change of the activity of the non-electrolyte produced by the salt, which has hitherto been the practice, but by the difference between the observed effect and the normal effect. In concentrated salt solutions the value of the latter given by (7) is not accurate, but nevertheless an approximate measure of the salting out may be obtained by the use of this value. Table II gives values of the quantity s, defined by

$$\log s = \Delta \log (p_2/p_2^0)/\Delta n_3 - 2/(n_1 + n_2),$$

for 1 m. lithium chloride solutions.

Table II.—Salting-out of alcohol from water-alcohol solutions by lithium chloride ($m = 1$).

N_2.	s	N_2.	s
0·02	1·23	0·25	1·12
0·05	1·19	0·50	1·11
0·10	1·16	0·70	1·08
0·15	1·14	0·90	1·04

Cases of negative salting out by certain salts which have been reported are apparently caused by neglect of the normal effect of the salt on the activity of the components of the solvent. For example, Sugden* found that while most salts increased the activity of acetic acid in dilute aqueous solution, its value in 1/N potassium nitrate was 0·998 that in water. He regarded this as an example of negative salting out; but since the theoretical lowering of the activity of the components of the solvent by a salt in 1/N solutions is 0·965, it is evident that the activity of acetic acid is in fact greater than if there were no salting out.

---

* 'J. Chem. Soc.,' vol. 128, p. 174 (1926).

The changes of the partial pressures of water and alcohol produced by the salt are related by (4), which may be written in the form

$$n_1 d\log(p_1/p_1^0)/dn_3 + n_2 d\log(p_2/p_2^0)/dn_3 = -0\cdot87. \tag{4A}$$

Table III gives the values of these quantities obtained from the smooth curves of fig. 1, for the interval 0-1·0 m. Exact agreement cannot be expected for this range of concentration, but the sum of the quantities on the left of (4A) is approximately constant and near to the theoretical value.

Table III.

N_2.	$n_2 d\log(p_2/p_2^0)/dn_3$.	$n_1 d\log(p_1/p_1^0)/dn_3$.	Sum.
0	0·0	−0·75	−0·75
0·25	+0·26	1·05	−0·79
0·50	+0·27	−1·12	−0·85
0·70	+0·06	−0·91	−0·85
0·80	−0·11	−0·65	−0·76
0·90	−0·39	−0·34	−0·73
1·00	−0·72	0·0	−0·72

The Free Energies of Lithium Chloride in Water-Alcohol Mixtures.

A direct determination of the free energy of transfer of lithium chloride from water to alcoholic solutions has not been feasible owing to the irregular behaviour of lithium amalgam electrodes in alcoholic solutions. However, if the variation of the free energy of two components of a series of ternary solutions is known, that of the third can be calculated by means of (3). Writing $\overline{F}_1 = \overline{F}_1^0 + RT\log p_1$, $\overline{F}_2 = \overline{F}_2^0 + RT\log p_2$, this takes the form

$$RT\,.\,n_1 d\log p_1 + RT\,.\,n_2\,d\log p_2 + n_3 d\overline{F}_3 = 0,$$

but since

$$n_1 d\log p_1^0 + n_2 d\log p_2^0 = 0,$$

where p_1^0, p_2^0 are the pressure of the corresponding solvent, we may make use of the more convenient expression

$$RT\,.\,n_1 d\log(p_1/p_1^0) + RT\,.\,n_2 d\log(p_2/p_2^0) + n_3 d\overline{F}_3 = 0.$$

This may be integrated for a constant value of n_3. The free energy of transfer of lithium chloride from its solution in water to that in any given solvent is then given by the integral:

$$\Delta\overline{F}_3 = -\frac{RT}{n_3}\int_{n_1^0}^{n'_1} n_1 d\log(p_1/p_1^0) - \frac{RT}{n_3}\int_0^{n'_1} n_2 d\log(p_2/p_2^0), \tag{8}$$

where n'_1 and n'_2 are the number of mols. of water and alcohol in the given solution containing the constant amount n_3 of lithium chloride, and n_1^0 is the quantity of water in the corresponding solution in water. These integrals have been evaluated graphically by plotting n_1 and n_2 against $\log (p_1/p_1^0)$ and $\log (p_2/p_2^0)$ respectively, and determining on squared paper the areas bounded by the smoothed curves to various points. The values obtained are given in calories in Table IV, and are shown plotted against the molar fraction of alcohol in fig. 2.

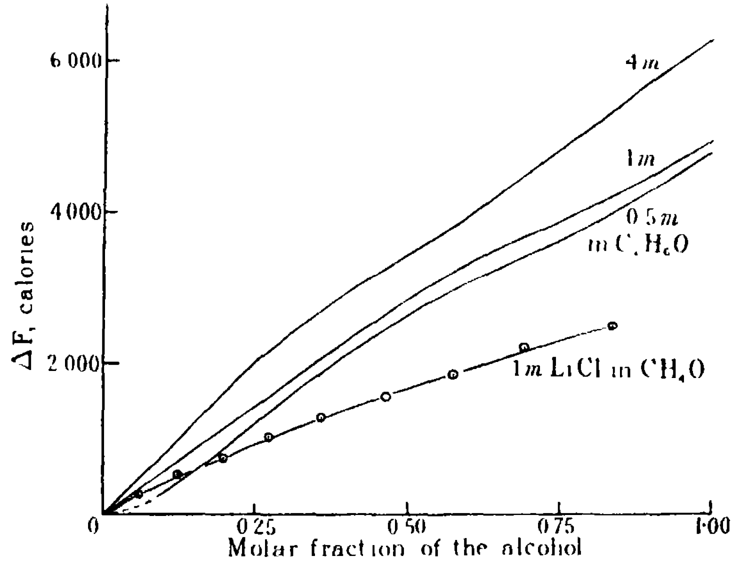


FIG. 2.—Free energies of transfer of lithium chloride from water to ethyl alcohol solutions (25°). —⊙— Åkerlöf's data for methyl alcohol solutions.

Table IV.—Free energies of transfer of lithium chloride from water to alcoholic solutions (calories) at 25°.

N_2.	$\Delta\bar{F}_v$.		
	0·5 m.	1·0 m	4·0 m.
0·1	250	560	750
0·2	840	1160	1620
0·3	1520	1740	2350
0·4	2140	2350	2950
0·5	2660	2900	3480
0·6	3090	3360	4000
0·7	3500	3760	4610
0·8	3910	4130	5210
0·9	4430	4580	5840
1·0	4860	5070	6410

An error in the experimental data at any point will affect all subsequent points in the process of integration. The probability of error is greatest for the 0·5 m. solutions containing small proportions of alcohol. The small vapour pressure lowerings of water in these solutions have not been determined with sufficient accuracy for this purpose, and it is probable that the first point on the 0·5 m. curve is too low. A correction of this point would affect the whole curve, which is otherwise nearly parallel with the curve for 1 m. solutions. Taking these facts into account, it is probable that the free energies of transfer of 0·5 m and 1·0 m. solutions are nearly the same.

The curves of $\Delta\overline{F}_3$ against N_2 are nearly linear, and it is probable that the divergence from linearity does not exceed the experimental error. This is in marked contrast to the behaviour of hydrogen chloride in these solvents,* where the greater part of the free-energy change occurs in high proportions of alcohol. Åkerlöf† has recently determined by electromotive measurements the free energies of transfer of the chlorides of lithium, sodium, potassium and hydrogen between water and methyl alcohol solutions and has found that whereas the alkali metal chlorides give an approximately linear relationship with the molar fraction of the alcohol, hydrogen chloride diverges widely in alcohol-rich solutions. His curve for lithium chloride ($m = 1$) is given for comparison in fig. 2. Scatchard‡ calculated from solubility data the free energies of transfer at infinite dilution of sodium and potassium chlorides from water to ethyl alcohol solutions and found a similar linear variation. Lithium chloride is too soluble to permit similar calculations from its solubility, but the calculations given above show that the behaviour of lithium chloride is approximately similar.

According to the simple electrostatic theory, regarding the ions as charged spheres, the free energy of transfer of a salt from a medium having the dielectric constant D_0 to a medium with dielectric constant D, at infinite dilution, is given by

$$\Delta\overline{F}_3{}^0 = -\Sigma \frac{e^2}{2r}(1/D_0 - 1/D), \tag{8}$$

where e is the charge and r the radius of an ion.§ Åkerlof derived from his measurements values of $\Delta\overline{F}_3$ at infinite dilution, and concluded that they were at least in approximate agreement with this equation. It is not possible to derive values for infinite dilution from our data, but it is of interest to compare the behaviour of lithium chloride at $m = 1$ in methyl and ethyl alcohols. Fig. 3 shows $\Delta\overline{F}_3$ for the two cases plotted against 1/D. Except in solutions containing a high proportion of methyl alcohol the slopes of the curves are approximately the same, showing that the free energies of transfer, at least

---

* Part I, 'Proc. Roy. Soc.,' A, vol. 125, p. 694 (1929).

† 'J. Amer. Chem. Soc.,' vol. 52, p. 2353 (1930).

‡ 'J. Amer. Chem. Soc.,' vol. 47, p. 2098 (1925).

§ We may observe that the free energy of transfer of lithium chloride between water and both methyl and ethyl alcohols is less than that of the chlorides of sodium and potassium requiring according to (8) a greater radius for the lithium ion. This might be ascribed to the greater degree of solvation of the lithium ion, but a satisfactory theory ought to include

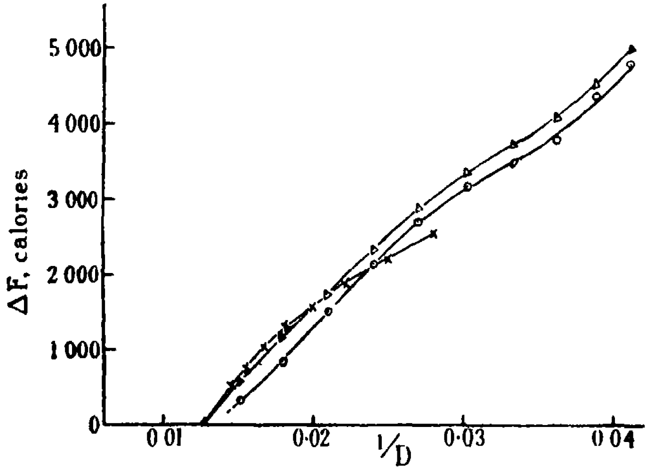


FIG. 3.—Variation of ΔF with 1/D. —△— LiCl, 1 m. in water-ethyl alcohol solutions. —⊙— LiCl, 0·5 *m* *in water-ethyl alcohol solutions.* —×— LiCl, 1 *m. in water*-methyl alcohol solutions.

up to 60 mols. per cent. alcohol, are to a first approximation dependent only on the dielectric constant of the medium and not on its nature.

the effect of solvation, at least so far as it is caused by electrical forces. Equation (8) cannot be regarded as satisfactory for it takes no account of the variation of the dielectric constant of the solvent in the vicinity of the ions. Webb (' J. Amer. Chem. Soc.,' vol. 48, p. 2589 (1926)) has calculated the electrical energy of charging spherical ions in water taking into account the electrical saturation of the solvent near the ions and has obtained values which differ considerably from those calculated by the simple formula $\Sigma e^2/2r \,.\, 1/D$, using the ordinary dielectric constant of water. It is thus evident that (8) cannot have more than a limited application. Although Åkerlöf concluded that his values of ΔF_3 for water-methyl alcohol solutions were in agreement with (8) within the errors of extrapolation, an examination of his curves shows that there is in fact a systematic divergence from the requirements of this equation, which becomes considerable in high concentrations of the alcohol. A more complete theory would have to take into account not only the electric saturation near the ions, but also the tendency of the more polarizable ions to congregate round the ions.

The senior author desires to express his appreciation of a Carnegie Teaching Fellowship, during the tenure of which most of this work was carried out.

Summary.

(1) The measurements of Shaw and Butler of the partial pressures of water and ethyl alcohol in solutions containing lithium chloride have been extended to solutions containing 2 and 4 mols per cent. of ethyl alcohol.

(2) A discussion is given of the thermodynamics of salting out. It is pointed out that the salting out is properly measured, not by the change of the activity of the non-electrolyte produced by the salt, but by the difference between this quantity and the normal effect, which is defined.

(3) The partial free energies of transfer of lithium chloride from water to alcoholic solutions have been calculated from the partial pressures of water and alcohol. The variation is approximately linear with the molar fraction of alcohol.

Contributions to the Mathematical Theory of Epidemics. III.—*Further Studies of the Problem of Endemicity.*

By W. O. KERMACK and A. G. MCKENDRICK.

(From the Laboratory of the Royal College of Physicians, Edinburgh.)

(Communicated by Sir Gilbert Walker, F.R.S.—Received January 27, 1933.)

Introduction.

In a previous paper[†] (Part II of this series) an attempt was made to treat from a general point of view the problem of a single disease in a population which consisted of three categories of people—namely, never infected, sick and recovered—and in which the infectivity of the disease was a function of the period of illness, whilst the susceptibility of a recovered person was a function of the period which had elapsed since the time of his recovery. New individuals entering the population either by birth or by immigration naturally entered the category of the never infected which for convenience we called "virgins." It was pointed out that the results obtained were subject to two important limitations: (1) that the disease under consideration was the only cause of death, and (2) that the age of the individuals did not affect their infectivity, susceptibility or reproductiveness.

[†] 'Proc. Roy. Soc.,' A, vol. 138, p. 55, (1932).

It is the purpose of the present paper to remove the first of these limitations by the introduction of constant non-specific death rates, which for the sake of generality are assumed to be different for virgins, sick, and recovered. It may be stated at once that the introduction of this additional factor produces surprisingly little change in the general nature of the results previously obtained, and that the conclusions of the previous paper hold with very little modification.

In the previous paper the results were first of all worked out for constant infectivity, recovery and death rates, and the more general problem, in which these rates were variable, was thereafter considered. The algebra for constant coefficients was relatively very simple. It is now found, however, that with the introduction of non-specific death rates, the simple case increases in complexity relatively much more than does the general case, so that the advantage of treating it separately largely disappears. In particular, whereas the various expressions for steady state levels previously came out explicitly as fairly simple functions of the constant coefficients, they are now dependent on a somewhat complicated quadratic equation. On the other hand the equations which refer to the case with variable rates, although they now contain a few extra terms, remain qualitatively similar in type to those previously obtained, and the same method of treatment leads at once to identical or closely similar results. We shall not therefore, in the present communication, treat in detail the case of constant rates but only give some of the main results. It may be mentioned, however, that the general equations have been checked at the various points by the introduction of constant rates, and comparison has been made between the formulæ so obtained and those found when constant rates were used throughout.

As in the previous paper the equations which describe the progress of small variations about the steady state are formulated, but their fuller discussion has at present been reserved.

It will be recalled that in Part II a number of special cases were discussed either because they had some special practical importance, or because they exhibited peculiarities from the mathematical point of view. With the introduction of non-specific death rates, a number of new special cases requiring detailed consideration came to light. To appreciate the relationship between all these cases it became necessary to adopt a scheme of classification, which although multiplying the total number of special cases considerably, made their treatment much simpler, and except in two instances only very brief discussion was required. The five special cases of Part II are readily accommodated in the new scheme.

A question of some practical importance is the effect upon the size of the population, the number of sick, and the relative incidence of the disease, of changes in the various parameters which characterize either the population or the disease. These points were investigated to some extent in the previous paper, but are more fully considered in the present communication both in connection with the general case, and with a number of the special cases referred to above. The results obtained are not always in accordance with expectation.

General Case.

General Equations.—A detailed description of the population under discussion will be found on pages 59 and 63 of the previous paper (Part II). The nomenclature and notation previously adopted remains unaltered, and to economize space will not be explained again here. The new death rates, which are now introduced, are denoted by $\bar{\pi}$, π and ρ for virgins, recovered persons, and sick, respectively, so that certain equations in Part II have to be modified. For convenience of reference the same numbers are used to denote the equations, an asterisk being added where alteration has been necessary In the case of equations which do not correspond to numbers in Part II, index numbers consecutive to those in that paper have been employed. It is to be noted that the variables employed refer throughout to population densities, but if the area be considered as fixed, the population size may be employed in place of the less usual conception of population density.

Equations (15), (16), and (17) become

$$\frac{d\bar{x}}{dt} + \overline{\pi x} = \bar{u}_t - \bar{v}_t. \tag{15*}$$

$$\frac{dx}{dt} + \pi x = u_t - \tilde{v}_t, \tag{16*}$$

$$\frac{dy}{dt} + \rho y = v_t - w_t - u_t, \tag{17*}$$

equations (19) to (27) remain unaltered.

To express $\bar{u}_{t\lambda}$ in terms of $\bar{u}_t$ we have the equation

$$\frac{\partial \bar{u}_{t\lambda}}{\partial t} + \frac{\partial \bar{u}_{t\lambda}}{\partial \lambda} = -\bar{u}_{t\lambda} \bar{f}(t) - \overline{\pi u}_{t\lambda}, \tag{28*}$$

whence

$$\bar{u}_{t\lambda} = \bar{u}_{t-\lambda}\, e^{-\int_0^{\lambda} [\bar{f}(t-\lambda') + \bar{\pi}]\, d\lambda'}. \tag{32*}$$

In a similar manner it can be shown that

$$v_{t\theta} = v_{t-\theta}\, e^{-\int_0^{\theta} (l_{\theta'} + d_{\theta'} + \rho)\, d\theta}, \tag{29*}$$

whence

$$\bar{f}(t) = \int_0^{\infty} \bar{k}_{\theta} v_{t-\theta} \mathrm{N}_{\theta}\, d\theta, \qquad \text{where} \qquad \mathrm{N}_{\theta} = e^{-\int_0^{\theta} (l_{\theta'} + d_{\theta'} + \rho)\, d\theta'} \tag{30*}$$

$$= \int_0^{\infty} \bar{\mathrm{K}}_{\theta} v_{t-\theta}\, d\theta, \qquad \text{where} \qquad \bar{\mathrm{K}}_{\theta} = \bar{k}_{\theta} \mathrm{N}_{\theta}. \tag{31}$$

By equation (19),

$$\bar{x} = \int_0^{\infty} \bar{u}_{t-\lambda}\, e^{-\int_0^{\lambda} [\bar{f}(t-\lambda') + \bar{\pi}]\, d\lambda'}\, d\lambda. \tag{33*}$$

Also

$$u_{t\tau} = u_{t-\tau}\, e^{-\int_0^{\tau} [f(t-\tau+\xi)\omega_{\xi} + \pi]\, d\xi}, \tag{37*}$$

$$= u_{t-\tau} \mathrm{F}(t-\tau, \tau), \quad \text{where} \quad \mathrm{F}(t-\tau, \tau) = e^{-\int_0^{\tau} [f(t-\tau+\xi)\omega_{\xi} + \pi]\, d\xi}, \tag{38*}$$

whence by equation (20)

$$x = \int_0^{\infty} \mathrm{F}(t-\tau, \tau)\, u_{t-\tau}\, d\tau. \tag{39}$$

Equations (40) to (45) remain formally unaltered.

Equations for Steady State.—In finding the conditions for a steady state certain modifications are necessary.

Clearly

$$\mathrm{L} + \mathrm{D} + \rho\mathrm{N} = \int_0^{\infty} (l_{\theta} + d_{\theta} + \rho)\, e^{-\int_0^{\theta} (l_{\theta'} + d_{\theta'} + \rho)\, d\theta'}\, d\theta = 1, \tag{46*}$$

while as before

$$\mathrm{U} = \bar{\mu}\bar{\mathrm{X}} + \mu\mathrm{X} + \nu\mathrm{Y} + m. \tag{47}$$

The other relations in equations (47), (48) and (49) no longer hold.

By (33*)

$$\bar{\mathrm{X}} = \bar{\mathrm{U}} \int_0^{\infty} e^{-\int_0^{\lambda} [\bar{f}(t-\lambda') + \bar{\pi}]\, d\lambda'}\, d\lambda,$$

but

$$\bar{f}(t) = \bar{\mathrm{K}}\mathrm{V}, \tag{50}$$

whence

$$\bar{\mathrm{X}} = \bar{\mathrm{U}} \int_0^{\infty} e^{-(\bar{\mathrm{K}}\mathrm{V} + \bar{\pi})\lambda}\, d\lambda$$

$$= \frac{\bar{\mathrm{U}}}{\bar{\mathrm{K}}\mathrm{V} + \bar{\pi}}. \tag{51*}$$

By (39)

$$\mathrm{X} = \mathrm{U}\int_0^\infty \mathrm{F}(t-\tau,\ \tau)\,d\tau.$$

But

$$\mathrm{F}(t-\tau,\ \tau) = e^{-\int_0^\tau [f(t-\tau+\xi)\,\omega_\xi+\pi]\,d\xi},$$

and

$$f(t) = \Phi\mathrm{V}, \tag{52}$$

hence

$$\int_0^\infty \mathrm{F}(t-\tau,\ \tau)\,d\tau = \int_0^\infty e^{-\int_0^\tau (\Phi\mathrm{V}\omega_\xi+\pi)\,d\zeta}\,d\tau,$$

$$= \mathrm{F}(\mathrm{V}), \tag{53*}$$

thus

$$\mathrm{X} = \mathrm{UF}(\mathrm{V}) \tag{54}$$

As previously

$$\mathrm{Y} = \mathrm{NV}, \qquad \mathrm{U} = \mathrm{LV} \qquad \text{and} \qquad \mathrm{X} = \mathrm{LVF}(\mathrm{V}). \tag{55)-(57}$$

Also

$$\tilde{\mathrm{V}} = \Phi\mathrm{VG}(\mathrm{V})\,\mathrm{U}, \tag{58) (in part}$$

where

$$\begin{aligned} \mathrm{G}(\mathrm{V}) &= \int_0^\infty \mathrm{G}(t-\tau,\ \tau)\,d\tau, \\ &= \int_0^\infty \omega_\tau\,\mathrm{F}(t-\tau,\ \tau)\,d\tau, \\ &= \int_0^\infty \omega_\tau\, e^{-\int_0^\tau (\Phi\mathrm{V}\omega_\zeta+\pi)\,d\zeta}\,d\tau, \\ &\frac{1-\pi\mathrm{F}}{\Phi\mathrm{V}}. \end{aligned} \tag{59*}$$

Thus

$$\tilde{\mathrm{V}} = \mathrm{U}(1-\pi\mathrm{F}). \tag{58*}$$

As before

$$\bar{\mathrm{V}} = \overline{\mathrm{XK}}\mathrm{V} \qquad \text{and} \qquad \mathrm{W} = \mathrm{DV}. \tag{60) and (61}$$

From these relations it follows that

$$\left.\begin{aligned} \mathrm{U} &= \mathrm{LV}, \\ \tilde{\mathrm{V}} &= \mathrm{LV}(1-\pi\mathrm{F}), \\ \bar{\mathrm{V}} &= \frac{\bar{\mathrm{K}}\mathrm{V}\bar{\mathrm{U}}}{\bar{\mathrm{K}}\mathrm{V}+\bar{\pi}}, \\ \mathrm{W} &= \mathrm{DV}, \\ \bar{\mathrm{X}} &= \frac{\bar{\mathrm{U}}}{\bar{\mathrm{K}}\mathrm{V}+\bar{\pi}}, \\ \mathrm{X} &= \mathrm{LFV}, \\ \mathrm{Y} &= \mathrm{NV}. \end{aligned}\right\} \tag{62*}$$

The condition for a steady state of the total population is by equations (15*), (16*) and (17*)

$$\bar{U} = W + \pi\bar{X} + \pi X + \rho Y,$$

whence after substituting the above values of $\bar{X}$, X, Y and W,

$$\bar{U} = \frac{(D + \pi LF + \rho N)(\bar{K}V + \bar{\pi})}{\bar{K}}.$$

But by equation (47),

$$\bar{U} = m + \bar{\mu}\bar{X} + \mu X + \nu Y,$$

whence, on substitution,

$$\bar{U} = \frac{(m + \mu LFV + \nu NV)(\bar{K}V + \bar{\pi})}{\bar{K}V + \bar{\pi} - \bar{\mu}}.$$

Continuing these two expressions for U, we have, for the steady state, V given by

$$\Theta(V) = LF\left\{\mu - \pi + \frac{\pi(\bar{\mu} - \bar{\pi})}{\bar{K}V}\right\} + \nu N - D - \rho N + \frac{m}{V} + \frac{(\bar{\mu} - \bar{\pi})(D + \rho N)}{\bar{K}V} = 0. \tag{65*}$$

Also the total number of individuals is

$$\begin{aligned} n &= \bar{X} + X + Y, \\ &= \frac{D + \pi LF + \rho N}{\bar{K}} + LFV + NV, \\ &= \frac{1 - L(1 - \pi F)}{\bar{K}} + (LF + N)V. \end{aligned} \tag{64*†}$$

We shall now consider the nature of the real positive roots of equation (65*), and we shall assume that $\bar{\mu} > \bar{\pi}$, and $\mu > \pi$.

Clearly as $V \to 0$, $\Theta(V) \to +\infty$. It is also obvious by inspection that $\Theta(V)$ decreases, as V increases. Thus the equation $\Theta(V) = 0$ will have one real positive root or no real positive roots, according as $\Theta(V)_{V\to\infty}$ is negative or positive.

Clearly

$$\Theta(V)_{V\to\infty} = \{LF(\mu - \pi) + \nu N - D - \rho N\}_{V\to\infty}.$$

---

† $\left[\text{Erratum: Equation (64) of Part II should read } n = \frac{D}{\bar{K}} + (LF + N)V\right]$.

As in Part II, p. 70, we shall assume that ω_ξ has the following properties. It is equal to zero between $\tau = 0$ and $\tau = \varepsilon$, it increases monotonically, and it has a constant value ω when $\tau > \eta$. Then

$$F = \int_0^\varepsilon e^{-\pi\tau}\,d\tau + \int_\varepsilon^\eta e^{-(\Phi V\bar{\Omega}_\tau + \pi\tau)}\,d\tau + \int_\eta^\infty e^{-\Phi V[\bar{\Omega}_\eta + \omega(\tau-\eta)] - \pi\tau}\,d\tau, \qquad (67^*)$$

$\left(\text{where as before } \bar{\Omega}_\tau = \int_0^\tau \omega_\xi\,d\xi\right)$,

$$= \frac{1 - e^{-\pi\varepsilon}}{\pi} + (\eta - \varepsilon)\,e^{-(\Phi V\bar{\Omega}_\sigma + \pi\sigma)} + \frac{e^{-(\Phi V\bar{\Omega}_\eta + \pi\eta)}}{\Phi V\omega + \pi},$$

(where σ is some value of τ between ε and η).

Thus

$$F(V)_{V\to\infty} = \frac{1 - e^{-\pi\varepsilon}}{\pi},$$

and

$$\Theta(V)_{V\to\infty} = (\mu - \pi)\,L\,\frac{(1 - e^{-\pi\varepsilon})}{\pi} + \nu N - D - \rho N. \qquad (68^*)$$

Therefore $\Theta(V) = 0$ has no real root if $D + \rho N - \nu N - (\mu - \pi)\,L\,\frac{(1 - e^{-\pi\varepsilon})}{\pi}$ is negative, and has one real root if this expression is positive. If it is zero then there is a root $V = \infty$.

Thus the sufficient conditions that there should be one and only one real positive root are

$$\left.\begin{aligned} &\bar{\mu} > \bar{\pi}, \qquad \mu > \pi \\ \text{and} \quad &D + \rho N - \nu N - (\mu - \pi)\,L\,\frac{(1 - e^{-\pi\varepsilon})}{\pi} > 0. \end{aligned}\right\} \qquad (103)$$

The third condition may also be written in the form

$$1 - L - \nu N - (\mu - \pi)\,L\,\frac{(1 - e^{-\pi\varepsilon})}{\pi} > 0.$$

It may be remarked that these conditions although sufficient are not necessary. When ε is zero the third condition becomes $\frac{D}{N} + \rho > \nu$, and in the case of constant coefficients this gives $d + \rho > \nu$. This means that the total deaths amongst the sick are greater than the births among the sick, and this is clearly necessary to balance the excess of births over deaths in the healthy, if a steady state is to be maintained.

As has been remarked in the introduction, we do not propose to give the detailed working for the case with constant coefficients. It may, however, be shown either by working from first principles, or by using equation (65*), calculating all the terms in it for constant coefficients, and substituting $V = Y(d + l + \rho)$, that Y is given by the quadratic

$$k\bar{k}(d+\rho-\nu)Y^2 - k\bar{k}\left\{m + \frac{(\bar{\mu}-\bar{\pi})(d+\rho)}{\bar{k}} - \frac{\pi(d+\rho-\nu)}{k} + (\mu-\pi)\frac{l}{k}\right\}Y$$
$$-\pi\{\bar{k}m + (\bar{\mu}-\bar{\pi})(d+l+\rho)\} = 0. \qquad (104)$$

The quadratic has one real positive and one real negative root if $\bar{\mu} > \bar{\pi}$, and $d + \rho > \nu$. On account of the complicated nature of the quadratic, it happens that it is in some respects easier to deal with the general case than with this very special case of constant coefficients.

Effect of Variations of Parameters on Steady State.—Having assumed that the three conditions found above are satisfied, we shall now investigate the variation of the absolute rate of incidence of the disease V, and of the proportional rate of incidence $T = \frac{V}{n}$, with changes in the immigration, birth and death rates.

From equation (65*) we have

$$\frac{\partial V}{\partial m} = -\frac{\dfrac{\partial\Theta}{\partial m}}{\dfrac{\partial\Theta}{\partial V}}, \qquad \frac{\partial V}{\partial\bar{\pi}} = -\frac{\dfrac{\partial\Theta}{\partial\bar{\pi}}}{\dfrac{\partial\Theta}{\partial V}}, \qquad \text{etc.} \qquad (105)$$

Also

$$\left.\begin{aligned} &\frac{\partial\Theta}{\partial m} = \frac{1}{V}, \qquad \frac{\partial\Theta}{\partial\bar{\mu}} = \frac{D + \rho N + L\pi F}{\bar{K}V} = \frac{1 - L(1-\pi F)}{\bar{K}V}, \\ &\frac{\partial\Theta}{\partial\mu} = LF, \qquad \frac{\partial\Theta}{\partial\nu} = N, \qquad \frac{\partial\Theta}{\partial\bar{\pi}} = -\frac{1 - L(1-\pi F)}{\bar{K}V}, \end{aligned}\right\} \qquad (106)$$

and

$$\frac{\partial\Theta}{\partial\pi} = -LF\frac{(\bar{K}V - \bar{\mu} + \bar{\pi})}{\bar{K}V} + L\left\{(\mu-\pi) + \pi\frac{(\bar{\mu}-\bar{\pi})}{\bar{K}V}\right\}\frac{\partial F}{\partial\pi}. \qquad (107)$$

But

$$\frac{\partial F}{\partial\pi} = -\int_0^\infty \pi e^{-\int_0^\tau(\Phi V\omega_\xi + \pi)\,d\xi}\,d\tau,$$

and is therefore negative; and, by (15*)

$$\bar{\pi}\bar{\bar{X}} = \bar{U} - \bar{V}$$
$$= m + \bar{\mu}\bar{\bar{X}} + \mu X + \nu Y - \bar{K}VX \qquad \text{(by (47) and (60))},$$

hence

$$\bar{X}(\bar{K}V - \bar{\mu} + \bar{\pi}) = m + \mu X + \nu Y,$$

which is a positive quantity, so that

$$\bar{K}V - \bar{\mu} + \bar{\pi} \tag{108}$$

is necessarily positive.

Thus $\partial\Theta/\partial m$, $\partial\Theta/\partial\bar{\mu}$, $\partial\Theta/\partial\mu$ and $\partial\Theta/\partial\nu$ are each positive, whilst $\partial\Theta/\partial\bar{\pi}$ and $\partial\Theta/\partial\pi$ are negative. As ρ is contained in the expressions N, L, D, F, etc., the value of $\partial\Theta/\partial\rho$ does not seem to be unambiguously either positive or negative.

Further

$$\frac{\partial\Theta}{\partial V} = LF'\left\{\mu - \pi + \frac{\pi(\bar{\mu} - \bar{\pi})}{\bar{K}V}\right\} - \frac{m}{V^2} - \frac{\{1 - L(1 - \pi F)\}(\bar{\mu} - \bar{\pi})}{\bar{K}V^2}, \tag{109}$$

and must be negative since F is positive, and F′ is negative.

Hence $\partial V/\partial m$, $\partial V/\partial\bar{\mu}$, $\partial V/\partial\mu$, and $\partial V/\partial\nu$ are each positive, whilst $\partial V/\partial\bar{\pi}$ and $\partial V/\partial\pi$ are negative.

[The sign of $\partial V/\partial\rho$ may be either positive or negative. It may readily be shown that if $\bar{\mu} = \bar{\pi}$ and $\mu = \pi$, $\partial\Theta/\partial\rho$ is negative, whilst if $\bar{\mu} > \bar{\pi}$, $\mu > \pi$, $\nu = 0$ and $l_\theta = 0$, $\partial\Theta/\partial\rho$ is positive]

Let $T = V/n$, so that, as before, 100T is the percentage rate of incidence of fresh cases of the disease, then

$$\frac{1}{T} = \frac{\bar{X}}{V} + \frac{X}{V} + \frac{Y}{V}$$

$$= \frac{1 - L(1 - \pi F)}{\bar{K}V} + LF + N, \tag{110}$$

and

$$-\frac{1}{T^2}\frac{\partial T}{\partial V} = \frac{-\{1 - L(1 - \pi F)\}}{\bar{K}V^2} + LF' + \frac{\pi LF'}{\bar{K}V}. \tag{111}$$

Since F′ is negative, $\partial T/\partial V$ is positive.

Consequently as T does not contain m, $\bar{\mu}$, μ, ν or $\bar{\pi}$ except implicitly in V, $\partial T/\partial m$, $\partial T/\partial\bar{\mu}$, $\partial T/\partial\mu$ and $\partial T/\partial\nu$ are positive, whilst $\partial T/\partial\bar{\pi}$ is negative.

The variation of T with π requires further consideration.

$$\frac{\partial T}{\partial\pi} = \frac{\partial' T}{\partial'\pi} + \frac{\partial T}{\partial V}\frac{\partial V}{\partial\pi},$$

where $\partial' T/\partial'\pi$ refers to partial differentiation with respect to π directly, and apart from V.

Thus

$$\frac{\partial \mathrm{T}}{\partial \pi} = -\frac{\mathrm{T}^2}{\dfrac{\partial \Theta}{\partial \mathrm{V}}} \begin{vmatrix} \dfrac{\partial' \frac{1}{\mathrm{T}}}{\partial' \pi}, & \dfrac{\partial \frac{1}{\mathrm{T}}}{\partial \mathrm{V}} \\[2ex] \dfrac{\partial \Theta}{\partial \pi}, & \dfrac{\partial \Theta}{\partial \mathrm{V}} \end{vmatrix}.$$

After substituting in the Jacobean and reducing, we find

$$\begin{aligned} \frac{\partial \mathrm{T}}{\partial \pi} = -\frac{\mathrm{T}^2}{\dfrac{\partial \Theta}{\partial \mathrm{V}}} \Bigg[\Bigg\{ & \frac{\mathrm{L}^2\mathrm{F}}{\bar{\mathrm{K}}\mathrm{V}} \frac{\partial \mathrm{F}}{\partial \mathrm{V}} \left(\mu - \pi + \pi \frac{\bar{\mu} - \bar{\pi}}{\bar{\mathrm{K}}\mathrm{V}} \right) \\ & - \frac{1 - \mathrm{L}\,(1 - \pi\mathrm{F})}{\bar{\mathrm{K}}\mathrm{V}^2}\,\mathrm{LF} + \frac{\mathrm{L}^2\,(\bar{\mathrm{K}}\mathrm{V} + \pi)}{\bar{\mathrm{K}}\mathrm{V}}\,\mathrm{F}\,\frac{\partial \mathrm{F}}{\partial \mathrm{V}}\,\frac{\bar{\mathrm{K}}\mathrm{V} - \bar{\mu} + \bar{\pi}}{\bar{\mathrm{K}}\mathrm{V}} \Bigg\} \\ & - \left(\frac{\mathrm{LF}}{\bar{\mathrm{K}}\mathrm{V}} + \mathrm{L}\,\frac{\bar{\mathrm{K}}\mathrm{V} + \pi}{\bar{\mathrm{K}}\mathrm{V}}\,\frac{\partial \mathrm{F}}{\partial \pi} \right) \frac{m}{\mathrm{V}^2} - \frac{1 - \mathrm{L}\,(1 - \pi\mathrm{F})}{\bar{\mathrm{K}}\mathrm{V}^2}\,\mathrm{L}\sigma\,\frac{\partial \mathrm{F}}{\partial \pi} \Bigg]. \end{aligned} \qquad (112)$$

The terms within the curly brackets are readily seen to be negative, whilst the last term is positive or negative according as

$$\sigma = (\bar{\mu} - \bar{\pi}) - (\mu - \pi), \qquad (113)$$

is greater or less than zero. Hence if $m = 0$, and $\sigma \leqslant 0$, $\partial \mathrm{T}/\partial \pi$ is negative. If $\sigma > 0$, or if m is not zero, the sign of $\partial \mathrm{T}/\partial \pi$ is not certain. By considering the extreme case $\omega = 0$, it may easily be shown that the coefficient of m is not necessarily negative.

So far we have been investigating the effect on the steady state of the disease of changes in the parameters which characterize the population. It is of interest to find the effects of alterations in the parameters which characterize the disease. These are $\bar{k}$, ϕ, ω, l, and d. Owing to the fact that l and d (like ρ) appear in N it does not seem possible to obtain any simple general result as to the effect of alteration of these parameters.

As the factor $\bar{k}$ enters into the equations for Θ and T only in the form of $\bar{\mathrm{K}}$, it is sufficient to investigate $\partial \mathrm{V}/\partial \bar{\mathrm{K}}$ and $\partial \mathrm{T}/\partial \bar{\mathrm{K}}$.

$$\frac{\partial \Theta}{\partial \bar{\mathrm{K}}} = -\frac{\{1 - \mathrm{L}\,(1 - \pi\mathrm{F})\}}{\bar{\mathrm{K}}^2\mathrm{V}}\,(\bar{\mu} - \bar{\pi}), \qquad (114)$$

and is necessarily negative [or zero if $\bar{\mu} = \bar{\pi}$], hence $\partial \mathrm{V}/\partial \bar{\mathrm{K}}$ is negative (or zero if $\bar{\mu} = \bar{\pi}$).

Now

$$\frac{\partial T}{\partial \bar{K}} = \frac{\partial' T}{\partial' \bar{K}} + \frac{\partial T}{\partial V}\frac{\partial V}{\partial \bar{K}},$$

$$= -\frac{T^2}{\dfrac{\partial \Theta}{\partial V}}\begin{vmatrix} \dfrac{\partial' \frac{1}{T}}{\partial \bar{K}}, & \dfrac{\partial \frac{1}{T}}{\partial V} \\ \dfrac{\partial \Theta}{\partial \bar{K}}, & \dfrac{\partial \Theta}{\partial V} \end{vmatrix}$$

$$= -\frac{T^2}{\dfrac{\partial \Theta}{\partial V}}\frac{1 - L(1 - \pi F)}{\bar{K}^2 V}\left(\frac{m}{V^2} + L\sigma\frac{\partial F}{\partial V}\right), \tag{115}$$

which is positive if $\sigma = (\bar{\mu} - \bar{\pi}) - (\mu - \pi)$ is negative or equal to zero. If σ is positive, the sign is ambiguous unless $m = 0$, in which case $\partial T/\partial K$ is negative. If $m = 0$ and $\sigma = 0$, $\partial T/\partial \bar{K} = 0$, that is, T is independent of $\bar{K}$.

Similarly as ϕ occurs in Θ and T only in the form of Φ it is sufficient to investigate $\partial V/\partial \Phi$ and $\partial T/\partial \Phi$.

$$\frac{\partial \Theta}{\partial \Phi} = \frac{\partial \Theta}{\partial F}\frac{\partial F}{\partial \Phi}$$

$$= L\{\mu - \pi + \pi(\bar{\mu} - \bar{\pi})\}\frac{\partial F}{\partial \Phi}. \tag{116}$$

Thus as $\partial F/\partial \Phi$ is negative, $\partial \Theta/\partial \Phi$ is negative, so that $\partial V/\partial \Phi$ is negative. Also

$$\frac{\partial T}{\partial \Phi} = -\frac{T^2}{\dfrac{\partial \Theta}{\partial V}}\begin{vmatrix} \dfrac{\partial' \frac{1}{T}}{\partial' \Phi}, & \dfrac{\partial \frac{1}{T}}{\partial V} \\ \dfrac{\partial \Theta}{\partial V}, & \dfrac{\partial \Theta}{\partial V} \end{vmatrix},$$

$$= -\frac{T^2}{\dfrac{\partial \Theta}{\partial V}}\frac{\delta F}{\partial \Phi}\begin{vmatrix} \dfrac{\partial' \frac{1}{T}}{\partial' F}, & \dfrac{\partial \frac{1}{T}}{\partial V} \\ \dfrac{\partial \Theta}{\partial F}, & \dfrac{\partial \Theta}{\partial V} \end{vmatrix},$$

$$= -\frac{T^2}{\dfrac{\partial \Theta}{\partial V}}\frac{\partial F}{\partial \Phi}\left(-\frac{1}{V^2}\right)\left[m\frac{\bar{K}V + \pi}{\bar{K}V} + \sigma\{1 - L(1 - \pi F)\}\right]. \tag{117}$$

Similarly it may readily be shown that $\partial V/\partial \omega$ is negative, and that $\partial T/\partial \omega$ is equal to the same expression as the above for $\partial T/\partial \Phi$, except that $\partial F/\partial \omega$

takes the place of $\partial F/\partial \Phi$, where $\partial F/\partial \omega$ like $\partial F/\partial \Phi$ is necessarily negative. It follows that $\partial T/\partial \Phi$ and $\partial T/\partial \omega$ are positive if σ is positive or equal to zero. If σ is negative $\partial T/\partial \Phi$ and $\partial T/\partial \omega$ are ambiguous in sign unless m is zero, in which case they are of negative sign. If m is zero and σ is zero, $\partial T/\partial \Phi$ and $\partial T/\partial \omega$ are both zero. It may be noted that if there is a general increase in infectivity or susceptibility, T will increase provided that σ is zero, unless m is also zero, when T will remain constant. If σ is not zero the net result would appear to be ambiguous.

From equations (65*), (110) and (46*) we may obtain T in the form

$$\left.\begin{aligned} T &= \frac{\bar{\mu} - \bar{\pi}}{1 - L + N(\bar{\mu} - \bar{\pi} - \nu) + \sigma LF - \frac{m}{\bar{V}}}, \\ \text{or} \quad T &= \frac{\bar{\mu} - \bar{\pi}}{D + N(\bar{\mu} - \bar{\pi} - \nu + \rho) + \sigma LF - \frac{m}{\bar{V}}} \end{aligned}\right\} \tag{118}$$

It is at once evident that if $m = 0$, and σ or $L = 0$, T is independent of $\bar{K}$, Φ and ω, thus verifying the above somewhat unexpected result. If in addition $\bar{\mu} - \bar{\pi} = \nu - \rho$, that is if the differences between birth and death rates are the same for all three categories—virgin, recovered and sick—then

$$T = \frac{\bar{\mu} - \bar{\pi}}{D}. \tag{119}$$

The conclusion that, if $\bar{\mu} - \bar{\pi} = \mu - \pi$ (a condition which, it might be expected will probably be satisfied in many cases), T is independent of the infectivity of the disease and of susceptibility to it, is one which would scarcely have been anticipated.

It is to be noticed that when V is zero, the equation

$$n = \frac{1 - L(1 - \pi F)}{\bar{K}} + (LF + N)V$$

becomes

$$n = \frac{1}{\bar{K}}. \tag{120}$$

In this case $\bar{X} = 1/\bar{K}$, $X = 0$, and $Y = 0$. It follows that no endemic disease can exist in a population which has a density of less than $1/\bar{K}$ This conclusion may be compared with the result found in the first paper of this series,† viz., that no epidemic can take place in a non-multiplying population provided that the density is less than $1/A$. If it be noted that $\bar{K}$ is essentially equivalent to A, it is clear that the two results in a general way agree with one another. Actually the *threshold* value $n_0 = \bar{X} = 1/\bar{K}$ can only exist as a steady state when $\bar{\mu} = \bar{\pi}$. If $\bar{\mu} = \bar{\pi} + \varepsilon$, where ε is a small quantity, and if $m = 0$, the following approximate results are readily found :—

---

† 'Proc. Roy. Soc.,' A, vol. 105, p. 700 (1927).

$$V = \frac{\varepsilon}{\bar{K}\left(1 - \nu N - L\frac{\mu}{\pi}\right)}, \tag{121}$$

$$Y = \frac{N\varepsilon}{\bar{K}\left(1 - \nu N - L\frac{\mu}{\pi}\right)}, \tag{122}$$

$$X = \frac{L\varepsilon}{\pi\bar{K}\left(1 - \nu N - L\frac{\mu}{\pi}\right)}, \tag{123}$$

$$X = \frac{1}{\bar{K}} - \frac{L\,\Phi\tilde{\Omega}\pi\varepsilon,}{\bar{K}^2\left(1 - \nu N - L\frac{\mu}{\pi}\right)}, \tag{124}$$

where

$$\tilde{\Omega} = \frac{1}{\pi}\int_0^\infty \omega_\tau e^{-\pi\tau}\,d\tau.$$

There are two possibilities :—

(1) $1 - \nu N - L\frac{\mu}{\pi} > 0$. In this case, when ε is small, V represents the unique positive solution known to exist, when the three conditions mentioned above are satisfied. When ε tends to 0, V tends to 0, and n tends to $1/\bar{K}$, so that $1/\bar{K}$ does in fact constitute a *threshold* value for n.

(2) $1 - \nu N - L\frac{\mu}{\pi} < 0$. In this case, when ε is small and positive, the real positive root which necessarily exists does not become zero with ε, so than $\varepsilon \to 0$ does not imply that $V \to 0$. It is, however, true that n is never less than $1/\bar{K}$, so that in this case $1/\bar{K}$ may be regarded as a lower limit to the density of a population in which there is an endemic disease in a steady state.

It is readily found that

$$X + \bar{X} = \frac{1}{\bar{K}} + \frac{L\varepsilon}{\bar{K}^2\left(1 - \nu N - L\frac{\mu}{\pi}\right)}\int_0^\infty (\bar{K} - \Phi\omega_\tau)\,e^{-\pi\tau}\,d\tau, \tag{125}$$

and therefore must exceed $1/\overline{\mathrm{K}}$, provided that $\Phi\omega_\tau$ never exceeds $\overline{\mathrm{K}}$, that is, as long as those recovered from the disease are not at any stage more liable to infection than the virgins.

Equations for Variations about Steady State.

The equations relating to the stability of the steady state are as follows :—

$$\frac{d\overline{x}'}{dt} = \overline{u}'_t - \overline{v}'_t - \overline{\pi x}', \tag{74*}$$

$$\frac{dx'}{dt} = u'_t - \tilde{v}'_t - \pi x', \tag{75*}$$

$$\frac{dy'}{dt} = v'_t - w'_t - u'_t - \rho y', \tag{76*}$$

$$\overline{u}'_t = \overline{\mu x}' + \mu x' + \nu y', \tag{77}$$

$$u'_t = \int_0^\beta \mathrm{L}_\theta v'_{t-\theta}\, d\theta, \qquad (\beta = \infty) \tag{78}$$

$$v'_t = \tilde{v}'_t + \overline{v}'_t, \tag{79}$$

$$\overline{v}'_t = \overline{x}'\overline{\mathrm{K}}\mathrm{V} + \overline{\mathrm{X}} \int_0^\beta \overline{\mathrm{K}}_\theta v'_{t-\theta}\, d\theta, \qquad (\beta = \infty) \tag{80}$$

$$w'_t = \int_0^\beta \mathrm{D}_\theta v'_{t-\theta}\, d\theta, \qquad (\beta = \infty) \tag{81}$$

$$y' = \int_0^\beta \mathrm{N}_\theta v'_{t-\theta}\, d\theta + \int_0^\beta v_{t\theta_1}\, d\theta_1, \qquad (\beta = \infty) \tag{82}$$

$$\overline{x}' = \int_0^\beta \overline{u}'_{t-\lambda} e^{-(\overline{\mathrm{K}}\mathrm{V}+\overline{\pi})\lambda}\, d\lambda - \overline{\mathrm{U}} \int_0^\beta e^{-(\overline{\mathrm{K}}\mathrm{V}+\overline{\pi})\lambda} \int_0^\lambda \int_0^{\beta-\lambda'} \overline{\mathrm{K}}_\theta v'_{t-\lambda'-\theta}\, d\theta\, d\lambda'\, d\lambda$$
$$+ \int_0^\infty \overline{u}_{t\lambda_1}\, d\lambda_1, \tag{84*}$$

$$\tilde{v}'_t = \Phi\mathrm{VU} \int_0^\beta \omega_\tau \mathrm{F}'(t-\tau, \tau)\, d\tau + \Phi\mathrm{V} \int_0^\beta \omega_\tau \mathrm{F}_\tau u'_{t-\tau}\, d\tau$$
$$+ \mathrm{UG} \int_0^\beta \Phi_\theta v'_{t-\theta}\, d\theta, \qquad (85)†$$

---

† In the previous paper a slight omission occurs in equation 85, which should read as above.

$$(\beta = \infty).\Bigg[\text{where} \qquad F_\tau = e^{-\int_0^\tau (\Phi V \omega_\xi + \pi)\, d\tau}$$

$$\text{and} \quad F'(t-\tau, \tau) = -F_\tau \int_0^\tau \int_0^\infty \Phi_\theta\, v'_{t-\tau+\xi-\theta}\, d\theta \omega_\xi\, d\xi,$$

$$\text{so that} \qquad F = \int_0^\infty F_\tau\, d\tau,$$

$$\text{and} \qquad G = \int_0^\infty \omega_\tau\, e^{-\int_0^\tau (\Phi V \omega_\xi + \pi)\, d\xi}\, d\tau.\Bigg]$$

and

$$x' = U \int_0^\beta F'(t-\tau, \tau)\, d\tau + \int_0^\beta F_\tau\, u'_{t-\tau}\, d\tau + \int_0^\beta u_{t\tau_1}\, d\tau_1 \,.\, (\beta = \infty). \tag{86}$$

Also we have

$$v_{t\theta_1} = {}_0v_{\theta_1}\, e^{\int_{\theta_1}^{t+T+\theta_1} (l_{\theta'} + d_{\theta'} + \rho)\, d\theta}, \tag{87*}$$

$$\bar{u}_{t\lambda_1} = {}_0\bar{u}_{\lambda_1}\, e^{-\int_{\lambda_1}^{t+T+\lambda_1} [\bar{f}(t-\lambda') + \bar{\pi}]\, d\lambda}, \tag{88*}$$

$$= {}_0\bar{u}_{\lambda_1}\, e^{-(\bar{K}V + \bar{\pi})(t+T)}, \tag{89*}$$

$$u_{t\tau_1} = {}_0u_{\tau_1} \frac{F(-T-\tau_1, t+T+\tau_1)}{F(-T-\tau_1, \tau_1)}, \tag{90}$$

$$= {}_0u_{\tau_1}\, e^{-\int_{\tau_1}^{t+T+\tau_1} (\Phi V \omega_\xi + \pi)\, d\xi} \tag{91*}$$

Equations (84*), (86), (82), (78), and (79) (using (85) and (80)) with upper limit $\beta = t + T$, correspond to the five equations (92) to (96) of the previous paper.

Special Cases.

So far the problem has been considered in its most general form. In the previous paper certain particular cases were discussed which appeared to be of special interest from the practical point of view. With the increased generality of the results now obtained, the number of special cases to be studied has naturally increased, and it becomes desirable to classify them according to some general scheme. It will be noted that in the theory as elaborated above, the conclusions arrived at depend upon certain relationships between $\bar{\mu}$, μ, ν, $\bar{\pi}$, π and ρ. It is therefore convenient in the first instance to classify the special cases according to the values of these six rates. We thus have the following five cases :—

Case 0 : general case, with conditions $\bar{\mu} \geqslant \bar{\pi}$, $\mu \geqslant \pi$, and

$$D + \rho N - \nu N - \frac{(1 - e^{-\pi c})}{\pi} L(\mu - \pi) > 0.$$

Case 1 : $\bar{\mu} = \mu$, $\bar{\pi} = \pi$, $\nu = 0$.

Case 2 : $\bar{\mu} = \bar{\pi}$, $\mu = \pi$ and $\nu = \rho$.

Case 3 : $\bar{\mu} = \mu = \nu = \bar{\pi} = \pi = \rho = 0$.

Case 4 : $\bar{\mu} = \mu = \nu = 0$.

It will be seen that the general conditions are satisfied in cases 0, 1, 2 and 3, and all the above theory holds. Case 4, however, is different in that the conditions $\bar{\mu} \geqslant \bar{\pi}$ and $\mu \geqslant \pi$ are definitely not satisfied, so that the question of the number of steady states in this instance requires special investigation.

If in the above classes we put $m = 0$, *i.e.*, exclude immigration, we have another set of special cases which we shall denote by 0′, 1′, 2′, etc Cases 0′ and 1′ involve no new features and no special comment is necessary, but case 2′ is characterized by the fact that the births balance the deaths from causes other than the disease under consideration, so that a steady state cannot exist unless $d_\theta = 0$. The same applies to case 3′. Further, case 4′ does not in any condition give a population in which a disease may exist in a steady state.

It is thus seen that, by considering the values of $\bar{\mu}$, μ, ν, $\bar{\pi}$, π, ρ and m, nine types are obtained of which only one requires special investigation from the point of view of the number of steady states which are possible.

It is, however, of special interest to subdivide each type according to the nature of the disease, which may be either fatal or non-fatal, and may or may not allow of recovery. Each type therefore gives rise to four varieties, namely:

(*a*) $l \neq$, $d \neq 0$,

(*b*) $l = 0$, $d \neq 0$,

(*c*) $l \neq 0$, $d = 0$,

(*d*) $l = 0$, $d = 0$.

These will be denoted by 0, 0_l, 0_d and 0_{ld}, etc. We have thus 36 different cases, but, as will be seen, certain of them do not give steady states corresponding to a finite population with endemic disease. For example, 3_d is obviously of this nature, as, by hypothesis, there are no deaths so that the population necessarily increases continually as the result of immigration. A steady state is therefore impossible. These cases will be indicated in the mathematical working by the emergence of values $V = 0$ or ∞.

The various cases are indicated in the following table, in which, for convenience of reference, the case numbers used in Part II are given in brackets. Where a blank occurs no steady state is possible.

	$l \neq 0, d \neq 0$	$l = 0, d \neq 0$	$l \neq 0, d = 0.$	$l = 0, d = 0.$
General case, $m \neq 0$	0	0_l	0_d	0_{ld}
General case, $m = 0$	$0'$	$0'_l$	$0'_d$	$0'_{ld}$
$\bar{\mu} = \mu, \bar{\pi} = \pi, \nu = 0, m \neq 0$	1	1_l	1_d	1_{ld}
$\bar{\mu} = \mu, \bar{\pi} = \pi, \nu = 0, m = 0$	$1'$	$1'_l$	$1'_d$	$1'_{ld}$
$\bar{\mu} = \bar{\pi}, \mu = \pi, \nu = \rho, m \neq 0$	2	2_l	—	—
$\bar{\mu} = \bar{\pi}, \mu = \pi, \nu = \rho, m = 0$	—	—	$2'_d$	$2'_{ld}$
$\bar{\mu} = \mu = \nu = \bar{\pi} = \pi = \rho = 0, m \neq 0$	3 (3)	3_l (4)	—	—
$\bar{\mu} = \mu = \nu = \bar{\pi} = \pi = \rho = 0, m = 0$	—	—	$3'_d$ (5)	$3'_{ld}$
$\left.\begin{matrix}\bar{\mu} = \mu = \nu = 0 \\ \bar{\pi} \neq 0, \pi \neq 0, \rho \neq 0\end{matrix}\right\} m \neq 0$	4	4_l	4_d	4_{ld}

Another contingency may arise, as in cases $2'_d$, $2'_{ld}$, $3'_d$ and $3'_{ld}$, in which the equation for V becomes indeterminate. In these cases the size of the population cannot alter, so that the total population is fixed by the initial conditions. A new equation for V has to be obtained by equating the sum of $\bar{X}$, X and Y to this fixed number n. The conditions under which this equation yields a definite and unique steady state require special investigation. It is sufficient to examine case $2'_d$ as the other three are particular cases of it It may be remarked that case $3'_d$ corresponds to case 5 in the previous paper. In the next two sections we shall therefore investigate the equations for V relating to cases 4 and $2'_d$. In addition a few remarks will be added in certain cases in which $l = 0$, a condition which makes the equation for V relatively simple, so that more definite results can be obtained than in the general case.

Case 4.—In case 4, $\bar{\mu} = \mu = \nu = 0$, and $m \neq 0$, that is to say there is immigration but no reproduction. It is obvious that the conditions $\bar{\mu} \geqslant \bar{\pi}$ and $\mu \geqslant \pi$ are not satisfied (except when $\bar{\pi}$ and π are both zero), so that the conclusions drawn above for the general case do not necessarily hold. It was pointed out above that the conditions for a definite and unique steady state were sufficient but not necessary, so that it is possible for definite and unique steady states to exist when these conditions are not satisfied. It happens, as will be shown below, that with certain assumptions the present special case is an example of such an exception.

The equation for V is in this case

$$V\Theta(V) = -LF\pi\frac{(\bar{K}V + \bar{\pi})}{\bar{K}} - (D + \rho N)V + m - \frac{\bar{\pi}}{\bar{K}}(D + \rho N) = 0, \quad (126)$$

hence when V is zero,

$$V\Theta(V) = -\frac{LF\pi\bar{\pi}}{\bar{K}} + m - \frac{\bar{\pi}}{\bar{K}}(D + \rho N). \tag{127}$$

But when V is zero

$$F = \int_0^\infty e^{-\pi\tau}\, d\tau = \frac{1}{\pi}, \tag{128}$$

hence in the limit $(V \to 0)$

$$V\Theta(V) = -\frac{L\bar{\pi}}{\bar{K}} + m - \frac{\bar{\pi}}{\bar{K}}(D + \rho N),$$

$$= m - \frac{\bar{\pi}}{\bar{K}}. \tag{129}$$

Also

$$\begin{aligned} \{V\Theta(V)\}_{V\to\infty} &= -LF\pi V - (D + \rho N)V, \\ &= -LV(1 - e^{-\pi s}) - (1 - L)V \text{ (above (68*))}, \\ &= -V(1 - Le^{-\pi s}), \end{aligned} \tag{130}$$

but L cannot be greater than unity so that $Le^{-\pi s} < 1$, hence

$$\{V\Theta(V)\}_{V\to\infty} \to -\infty. \tag{131}$$

Thus there will certainly be one real root, and there may be an odd number of real roots if $m > \bar{\pi}/\bar{K}$, whilst there may be no real roots or an even number if $m < \bar{\pi}/\bar{K}$.

We shall now find the conditions which will ensure that $V\Theta(V)$ decreases monotonically with V.

$$-V\Theta(V) = (1 - L)V + LF\pi\frac{\bar{K}V + \bar{\pi}}{\bar{K}} - m + \frac{\bar{\pi}}{\bar{K}}(1 - L).$$

Now

$$(\bar{K}V + \bar{\pi})F = (\bar{K}V + \pi)F - (\pi - \bar{\pi})F,$$

but

$$\begin{aligned} (\bar{K}V + \pi)F &= \int_0^\infty (\bar{K}V + \pi)\, e^{-\int_0^\tau (\Phi V\omega_\xi + \pi)\, d\xi}\, d\tau \\ &= \int_0^\infty (\Phi V\omega_\tau + \pi)\, e^{-\int_0^\tau (\Phi V\omega_\xi + \pi)\, d\xi}\, d\tau \\ &\qquad + \int_0^\infty V(\bar{K} - \Phi\omega_\tau)\, e^{\int_0^\tau (\Phi V\omega_\xi + \pi)\, d\xi}\, d\tau \\ &= 1 - \int_0^\infty \frac{\bar{K} - \Phi\omega_\tau}{\Phi\omega_\tau} e^{\pi\tau}\, de^{-\int_0^\tau \Phi V\omega_\xi d\xi} \\ &= 1 + \left(\frac{\bar{K}}{\Phi\omega_0} - 1\right) - \int_0^\infty \left\{\pi\left(\frac{\bar{K}}{\Phi\omega_\tau} - 1\right) + \frac{\bar{K}}{\Phi\omega^2_\tau}\frac{d\omega_\tau}{d\tau}\right\} e^{-\int_0^\tau (\Phi V\omega_\xi + \pi)\, d\xi}\, d\tau, \end{aligned} \tag{132}$$

where ω_0 is the value of ω_τ when τ is zero. We shall temporarily assume that ω_0 is not zero. We shall also assume as in the previous paper that $d\omega_\tau/d\tau$ is never negative, and further that $\Phi\omega_\tau$ is never greater than $\bar{K}$; that is to say, that the infectivity of the recovered never decreases, and that it never exceeds the infectivity of the virgins. If these conditions are satisfied the expression

$$\pi\left(\frac{\bar{K}}{\Phi\omega_\tau} - 1\right) + \frac{\bar{K}}{\Phi\omega^2_\tau}\frac{d\omega}{d\tau}$$

is never negative.

We have in the above assumed that ω_0 is not zero. It may, however, easily be shown by a method similar to that employed in the previous paper (p. 79) that this assumption is not essential.

If, further, π is not less than $\bar{\pi}$, the term $-(\pi - \bar{\pi})F$ will increase with V (or remain constantly zero) so that $F(\bar{K}V + \bar{\pi})$ will increase with V. Also the term $(1 - L)V$ cannot decrease with V, since $L \leqslant 1$, so that $-V\Theta(V)$ increases with V, or $V\Theta(V)$ decreases with V provided that

$$\pi \geqslant \bar{\pi}, \tag{133}$$

$$\omega_\tau \leqslant \frac{\bar{K}}{\Phi} \text{ for all values of } \tau, \tag{134}$$

and

$$d\omega/d\tau \tag{135}$$

is never negative.

If these three conditions are satisfied a unique steady state will exist provided that $m > \bar{\pi}/\bar{K}$, whilst no steady state in which the disease is endemic will exist if $m < \bar{\pi}/\bar{K}$. Conditions (133) to (135) imply that the chances of death of the members of the population from general causes are not decreased by an attack of the disease; that the susceptibility after an attack remains steady or increases monotonically but never exceeds that which exists in the virgin individual.

We shall now investigate the dependence of V upon m, $\bar{\pi}$, π and ρ. As $V\Theta(V)$ constantly decreases with V, provided that the above conditions are satisfied,

$$\frac{\partial V\Theta(V)}{\partial V}$$

is always negative. Also

$$\left.\begin{aligned} &\frac{\partial}{\partial m}V\Theta(V) = 1\,;\ \frac{\partial}{\partial\bar{\pi}}V\Theta(V) = -\frac{LF\pi}{\bar{K}} - \frac{(1 - L)}{\bar{K}} \\ &\text{and is therefore negative;} \\ &\frac{\partial}{\partial\pi}V\Theta(V) = -L\frac{(\bar{K}V + \bar{\pi})}{\bar{K}}\left(F + \pi\frac{\partial F}{\partial\pi}\right). \end{aligned}\right\} \tag{136}$$

But

$$\pi \frac{\partial \mathrm{F}}{\partial \pi} = -\int_0^\infty \pi\tau e^{-\int_0^\tau \Phi \mathrm{V} \omega_\xi d\xi} e^{-\pi\tau} d\tau$$

$$= \Big| \tau e^{-\int_0^\tau \Phi \mathrm{V} \omega_\xi d\xi} e^{-\pi\tau} \Big|_0^\infty - \int_0^\infty e^{-\pi\tau} e^{-\int_0^\tau \Phi \mathrm{V} \omega_\xi d\xi} d\tau$$

$$+ \int_0^\infty \tau \Phi \mathrm{V} \omega_\tau e^{-\int_0^\tau \Phi \mathrm{V} \omega_\xi d\xi} e^{-\pi\tau} d\tau \tag{137}$$

$$= -\mathrm{F} + \text{a positive quantity.}$$

Hence $\mathrm{F} + \pi \frac{\partial \mathrm{F}}{\partial \pi}$ is positive, and consequently $\frac{\partial}{\partial \pi} \mathrm{V}\Theta(\mathrm{V})$ is negative.

Further

$$\frac{\partial \mathrm{V}\Theta(\mathrm{V})}{\partial \rho} = \frac{\partial \mathrm{V}\Theta(\mathrm{V})}{\partial \overline{\mathrm{K}}} \frac{\partial \overline{\mathrm{K}}}{\partial \rho} + \frac{\partial \mathrm{V}\Theta(\mathrm{V})}{\partial \mathrm{L}} \frac{\partial \mathrm{L}}{\partial \rho} + \frac{\partial \mathrm{V}\Theta(\mathrm{V})}{\partial \mathrm{F}} \frac{\partial \mathrm{F}}{\partial \rho}.$$

Now

$$\frac{\partial \mathrm{V}\Theta(\mathrm{V})}{\partial \overline{\mathrm{K}}} = \frac{\mathrm{LF}\pi\overline{\pi}}{\overline{\mathrm{K}}^2} + \frac{(1-\mathrm{L})\,\overline{\pi}}{\overline{\mathrm{K}}^2}, \tag{138}$$

which is positive.

Also

$$\frac{\partial \mathrm{V}\Theta(\mathrm{V})}{\partial \mathrm{L}} = \left(\mathrm{V} + \frac{\overline{\pi}}{\overline{\mathrm{K}}}\right)(1 - \pi\mathrm{F}); \tag{139}$$

but

$$\mathrm{F} = \int_0^\infty e^{(\pi\tau + \int_0^\tau \Phi \mathrm{V} \omega_\xi d\xi)} d\tau.$$

$$\leqq \int_0^\infty e^{-\pi\tau} d\tau,$$

$$\leqq \frac{1}{\pi};$$

hence $1 - \pi\mathrm{F} \geqq 0$, and $\frac{\partial \mathrm{V}\Theta(\mathrm{V})}{\partial \mathrm{L}}$ is positive (or zero in the particular case where ω_τ is always zero).

Again

$$\frac{\partial \mathrm{V}\Theta(\mathrm{V})}{\partial \mathrm{F}} = -\mathrm{L}\pi \frac{\overline{\mathrm{K}}\mathrm{V} + \overline{\pi}}{\overline{\mathrm{K}}}, \tag{140}$$

which is negative.

As $\partial \overline{\mathrm{K}}/\partial \rho$ and $\partial \mathrm{L}/\partial \rho$ are negative, whilst $\partial \mathrm{F}/\partial \rho$ is zero, $\partial \mathrm{V}\Theta(\mathrm{V})/\partial \rho$ is always negative.

It follows from the above that $\partial V/\partial m$ is always positive, whilst $\partial V/\partial \bar{\pi}$, $\partial V/\partial \pi$ and $\partial V/\partial \rho$ are negative provided that the conditions (133) to (135) are satisfied. Also since $\partial T/\partial V$ is positive (111), it follows that $\partial T/\partial m$ is positive, and $\partial T/\partial \bar{\pi}$ is negative. The signs of $\partial T/\partial \pi$ nas $\partial T/\partial \rho$ appear to be ambiguous.

In this case then, when the conditions (133) to (135) are satisfied and $m > \bar{\pi}/\bar{K}$, the absolute rate of infection V, and the proportional rate T both increase with the rate of immigration m, but decrease with increase in the general death rate $(\bar{\pi})$ of the virgin group If $m < \bar{\pi}/\bar{K}$ no disease will exist in the community, so that $m = \bar{\pi}/\bar{K}$ is the threshold immigration rate.

As regards the dependence of V upon $\bar{K}$, Φ and ω, we find

$$\frac{\partial V\Theta(V)}{\partial \bar{K}} = \frac{\bar{\pi}}{\bar{K}^2}\{1 - L(1 - \pi F)\}, \tag{141}$$

$$\frac{\partial V\Theta(V)}{\partial \Phi} = -L\frac{\bar{\pi}}{\bar{K}}(\bar{K}V + \bar{\pi})\frac{\partial F}{\partial \Phi}, \tag{142}$$

and

$$\frac{\partial V\Theta(V)}{\partial \omega} = -L\frac{\bar{\pi}}{\bar{K}}(\bar{K}V + \bar{\pi})\frac{\partial F}{\partial \omega}, \tag{143}$$

which are each positive, hence

$$\frac{\partial V}{\partial \bar{K}}, \quad \frac{\partial V}{\partial \Phi} \quad \text{and} \quad \frac{\partial V}{\partial \omega}$$

are positive.

Also

$$\frac{\partial T}{\partial \bar{K}} = \frac{-T^2}{\dfrac{\partial V\Theta(V)}{\partial V}}\frac{\{1 - L(1 - \pi F)\}}{\bar{K}^2 V}\left\{\frac{1 - L(1 - \pi F)}{\bar{K}V}(\bar{K}V + \pi) - L\pi\frac{\partial F}{\partial V}\frac{\sigma}{\bar{K}}\right\} \tag{144}$$

where σ reduces in this case to $\pi - \bar{\pi}$, and

$$\frac{\partial T}{\partial \Phi} = \frac{-T^2}{\dfrac{\partial V\Theta(V)}{\partial V}}\left(-\frac{L}{\bar{K}V}\frac{\partial F}{\partial \Phi}\right)\left[\{1 - L(1 - \pi F)\}(\bar{K}V + \pi) + \pi\frac{(\bar{K}V + \bar{\pi})\{1 - L(1 - \pi F)\}}{\bar{K}V}\right]. \tag{145}$$

$\partial T/\partial \omega$ is equal to the same expression as the above for $\partial T/\partial \Phi$ except that $\partial F/\partial \omega$ takes the place of $\partial F/\partial \Phi$, where $\partial F/\partial \omega$ like $\partial F/\partial \Phi$ is necessarily negative.

It follows that $\partial T/\partial \Phi$ and $\partial T/\partial \omega$ are positive, whilst $\partial T/\partial \bar{K}$ is positive if σ is positive or zero, and ambiguous when σ is negative ($m \neq 0$).

In the special case 4_l where $l_0 = 0$, that is to say where there is no recovery from the disease, we have $V = m - \bar{\pi}/\bar{K}$, and all the quantities $\bar{X}$, X, etc., are given explicitly in terms of the various constants. The results are in general the same as in case 4, except that the conditions (133) to (135) are now meaningless, as the constants concerned do not enter into the problem.

Case $2'_d$.—This case which calls for special consideration refers to a community in which there is a non-fatal disease, but which does not form a completely closed system. Births and deaths from other causes are taking place, but at such a rate as to balance each other, so that the total number of the population remains constant. Immigration is excluded. The conditions are $m = 0$, $\bar{\mu} = \bar{\pi}$, $\mu = \pi$, $\nu = \rho$ and $d = 0$ (whence $L + \rho N = 1$). The equation for V vanishes completely, and we find that if $n = \bar{X} + X + Y$, $dn/dt = 0$. V is thus defined by the equation $n = \bar{X} + X + Y$ where n is the total number of individuals in the system. Thus

$$n = \frac{1 - L(1 - \pi F)}{\bar{K}} + LVF + NV. \tag{147}$$

We have to examine the nature of the roots of this equation. Let

$$\chi(V) = \frac{L}{\bar{K}}(\bar{K}V + \pi)F + \frac{1 - L}{\bar{K}} + NV - n \tag{148}$$

It is readily seen that

$$\chi(V)_{V \to \infty} \to +\infty,$$

also

$$\chi(V)_{V \to 0} \to \frac{L\pi}{\bar{K}} F_{V \to 0} + \frac{1 - L}{\bar{K}} - n$$

But

$$F_{V \to 0} \to \int_0^\infty e^{-\pi\tau}\, d\tau = 1/\pi,$$

as by assumption ω_τ is always finite.

Hence

$$\chi(V)_{V \to 0} \to \frac{L}{\bar{K}} + \frac{1 - L}{\bar{K}} - n = \frac{1}{\bar{K}} - n. \tag{149}$$

There may thus be an even number of roots if $n < 1/\bar{K}$, and there will be an odd number if $n > 1/\bar{K}$. Clearly $n_0 = 1/\bar{K}$ represents a *threshold* density in the same sense as $n_0 = 1/\Phi\omega$ represented a *threshold* density in case 5 of the previous paper (p. 79), *i.e.*, $3'_d$ according to the present system of classification.

We shall now find the conditions under which $\chi(\mathrm{V})$ increases monotonically with V, so that there will be one root or none at all according as n is greater or less than $1/\bar{\mathrm{K}}$

We have shown above (132) that $(\bar{\mathrm{K}}\mathrm{V} + \pi)\,\mathrm{F}$ is never negative provided that $d\omega_\tau/d\tau$ is never negative, and $\Phi\omega_\tau$ is never greater than $\bar{\mathrm{K}}$, that is to say that the infectivity of the recovered never decreases, and that it never exceeds the infectivity of the virgins. We have in fact

$$\chi(\mathrm{V}) = 1 + \left(\frac{\bar{\mathrm{K}}}{\Phi\omega_0} - 1\right) - \int_0^{\infty}\left\{\pi\left(\frac{\bar{\mathrm{K}}}{\Phi\omega_\tau} - 1\right) + \frac{\bar{\mathrm{K}}}{\Phi\omega^2}\frac{d\omega}{d\tau}\right\} e^{-\int_0^{\tau}(\Phi\mathrm{V}\omega_\xi)\,d\xi}\, e^{-n\tau}\, d\tau$$

$$+ \frac{(1 - \mathrm{L})}{\bar{\mathrm{K}}} + \mathrm{NV} - n, \qquad (150)$$

and, as the expression within curly brackets within the integral is never negative on the above assumptions, it is readily seen that $\chi(\mathrm{V})$ always increases with V. Thus the equation $\chi(\mathrm{V}) = 0$ has one and only one real positive root provided that $n > 1/\bar{\mathrm{K}}$, whilst if $n < 1/\bar{\mathrm{K}}$ there is no real positive root. When $n = 1/\bar{\mathrm{K}}$ there is only one root namely $\mathrm{V} = 0$

In case $2'_{ld}$, *i.e.*, an incurable but non-fatal disease,

$$\mathrm{V} = \rho\left(n - \frac{1}{\mathrm{K}}\right). \qquad (151)$$

The more general conclusions arrived at above hold, except that the condition $\bar{\mathrm{K}} \geqslant \Phi\omega_\tau$ is now meaningless, and therefore unnecessary. The *threshold* density $n_0 = 1/\bar{\mathrm{K}}$ exists as before.

Other Special Cases.—As when $l_\theta = 0$, V can always be expressed explicitly in terms of the various constants, it is desirable to consider the most general case of this type, namely 0_l. We have

$$\mathrm{V} = \frac{m + \dfrac{\bar{\mu} - \bar{\pi}}{\mathrm{K}}}{1 - \nu\mathrm{N}}. \qquad (152)$$

This is certainly positive provided that $\bar{\mu} > \bar{\pi}$, and $\nu\mathrm{N} < 1$. The latter condition is the third condition for the existence of a unique steady state (103), when L is zero. The second condition, namely $\mu > \pi$, becomes meaningless, and is therefore unnecessary.

It is still impossible to determine unambiguously the sign of $\partial V/\partial \rho$, but if $\nu = 0$, it may readily be shown that $\partial V/\partial \rho$ is positive, and $\partial S/\partial \rho$ is negative, where $S = \dfrac{Y}{X + Y}$, *i.e.*, the proportion of sick These results may be verified for constant coefficients, and it may also be shown that in this case the equilibrium is a stable one.

A few remarks may be added regarding a special group of cases, which are important in practice but do not require special treatment as no fundamentally new consideration arises. When permanent and complete immunity is conferred by an attack of the disease, $\omega = 0$, and $F = 1/\pi$ The equation for V becomes

$$V = \frac{m + \dfrac{\bar{\mu} - \bar{\pi}}{\bar{K}}}{1 - \nu N - L\dfrac{\mu}{\pi}}. \tag{153}$$

These cases are closely related to the group in which $l_\theta = 0$ and are probably of considerable importance as describing the group of virus diseases in which immunity appears often to be permanent and complete.

Discussion.

The main results detailed in the above investigation may be summarized by saying that the existence of a death rate from causes other than the particular disease which is operating, does not materially alter any of the results obtained in Part II of this investigation We have not yet taken into account the effect of variation in age-constitution, nor have variations in natural immunity been allowed for.

The chief feature of the systems treated here, and in the previous paper, is the existence of a steady state, which is unique provided that certain conditions, which are likely to exist in nature, are satisfied (equation (103), also (133) to (135)). This steady state is naturally a function both of the parameters characteristic of the disease—the infectivity, death and recovery rates—and of the other parameters which refer more particularly to the population—the birth rate, the immigration rate, and the non-specific death rates.

The incidence of disease is raised when the immigration or birth rates are increased (p. 102). This statement holds whether the incidence of disease is measured by (1) the number of sick Y; (2) the relative number of sick $S = Y/n$; (3) the incidence rate V; or (4) the relative incidence rate

$T = V/n$. Further, the incidence constantly decreases with increase in the death rate of the virgins, but whereas Y and V decrease with increase in the death rate of the recovered, the behaviour of S and T is more complex and depends upon the value of σ (112). The effect of the non-specific death rate of the sick (ρ) is complicated, and obscure.

When we turn to the effect on the incidence of the characteristic features of the disease (p. 103), the most interesting results are those relating to the effect of variations in infectivity. It is found that by such an alteration both the total population and the number of diseased persons in it are changed For example, increase in the susceptibility of the virgins, leading to an increase in $\bar{K}$, causes a decrease in the size of the population n, and also a decrease both of Y and of V. The effect on T and S is more complicated and depends upon the value of σ (113). If there is no immigration and $\sigma = 0$, then change in susceptibility does not alter T or S (p. 105). Likewise Y and V both decrease with increase in ω_r, the susceptibility of the recovered, as well as with Φ which measures the infectivity of the disease ; whilst the changes undergone by T and S again depend upon the value of σ. As before, when $m = 0$, and $\sigma = 0$, T and S do not alter with changes in susceptibility or infectivity. It follows that an increase in infectivity resulting in greater chances of infection of both virgins and recovered will decrease both V and Y, but will not alter T and S provided that $m = 0$ and $\sigma = 0$. If, however, these latter conditions are not satisfied, it does not appear possible to predict in a general way what alteration in T and S will ensue. It is to be noted that the condition $\sigma = 0$ means that the difference between the birth rate and the non-specific death rate is the same for virgins as for recovered, a condition which is usually approximately fulfilled. It is at first sight surprising that in these circumstances ($m = 0$ and $\sigma = 0$), alterations in infectivity or susceptibility should be without effect on the relative prevalence of the disease in the community, but the fact is that the relative prevalence can then be expressed by a formula which does not involve either $\bar{\bar{K}}$ Φ, or ω_r (equation (118)) so that changes in these parameters do not affect the result.

It will be seen from the above that, other things being equal, decrease in infectivity or susceptibility is always an advantage since it enables a larger total population density to exist. If, however, the goal desired is to reduce the relative disease rate in the community to a minimum, then it is far from certain that decrease in infectivity or susceptibility of the disease will bring about the desired result, whilst it would appear certain that the absolute number of the diseased, as well as the absolute rate of incidence, will actually increase as the infectivity or susceptibility falls. There seems, of course, little doubt that the immediate effect of a reduction in infectivity or susceptibility is to bring about a fall in all these numbers, but the population density then begins to increase and when a steady state is again reached the increase is of such a magnitude as to more than compensate for the immediate effects It must be emphasized that, when the effect of the variation of a parameter is considered in the present section, the comparison is between the steady state which is possible after the parameter has been varied, and that which existed before the variation took place. We are not concerned with the immediate effect of the parameter variation on the system, which will temporarily have been disturbed from a steady state.

It has been shown (p. 105) in the general case that when disease exists in a community the population density is necessarily greater than $1/\bar{K}$. The existence of a *threshold* value in relation to an observed quantity implies that the process under consideration (*e.g*, the existence of disease in a steady state) may exist whenever the quantity in question exceeds that value, whilst it cannot exist when the quantity is equal to or less than that value. Let us consider a system in which the quantity in question is initially below the threshold The system may be such that the quantity cannot change its value, and therefore the process or event to which the threshold refers, can never under any circumstances occur. A threshold of this nature may be called a *threshold of the first type.* On the other hand the system may be such that, when the magnitude in question is below the threshold, the system gradually alters until the threshold is exceeded. The process to which the threshold has reference may then appear, but it could not possibly have appeared until this adjustment had taken place That the magnitude should initially exceed the threshold value is not then necessary in order that the process should ultimately exhibit itself, nevertheless the process will not exhibit itself until the magnitude has exceeded its threshold value. Such a threshold may be called a *threshold of the second type.* It is clear that the population density $1/\bar{K}$ referred to above is a threshold of the second type, whilst the threshold in case 4 (p 114) referring to immigration, and that in case $2'_d$ (p. 115) referring to the total population, are of the first type.

In case $2'_d$ the population though not really a closed one is virtually of this nature, because the births just balance the non-specific deaths, and here a threshold value $(1/\bar{K})$ of the total population exists which is such that no disease can occur if the population density is less than this quantity. This result may be compared with that obtained in case 5 of the second paper of this series, that is to say in case $3'_a$ in the present classification, where a threshold value of $1/\Phi\omega$ was found. This latter case is obtained from case $2'_a$ by putting $\bar{\mu}$, μ, ν, $\bar{\pi}$, π and ρ equal to zero, so that one would at first sight expect that the threshold would continue to be $1/\bar{K}$. However, as there are no births, $\bar{K}$, which refers to the virgins, comes ultimately to have no significance, and the result shows that under these conditions the threshold alters to $1/\Phi\omega$, which by hypothesis, is either greater than or equal to $1/\bar{K}$.

In case 4 it is the rate of immigration m which has a threshold value If the immigration rate fails to reach the value $\bar{\pi}/\bar{K}$ then the disease cannot exist in the community in an endemic form. If $\bar{\pi}$ is very small, the threshold value will also tend to be very low, that is to say, very slow immigration would keep the disease going, but it would appear that if, as in most cases, non-specific deaths cannot be absolutely excluded, and if the population density is sufficiently great, then the threshold immigration rate, though small, will be finite.

In a series of interesting papers Greenwood and Topley and their collaborators* have investigated the progress of disease in communities of mice under various experimental conditions. Amongst their conclusions they emphasize the influence on the progress of the infection of immigration of healthy animals into the community For example† "The sole condition required for the indefinite propagation of an endemic disease is a continuous immigration of susceptible individuals." The conditions of the experiments of these authors would seem to be approximately represented by case 4 above, as the mice were not increasing in number by reproduction, and at the same time were subject to a certain number of non-specific deaths. Further they come to the conclusion that the increase in resistance of exposed animals as compared with that of fresh individuals, is caused by active immunization as the result of an attack of the disease, rather than to selection working upon individual differences originally present This is accommodated by the condition that $\Phi\omega$ is less than $\bar{K}$, which we found it necessary to assume, p. 112 It would therefore be expected that a threshold rate of immigration would exist under these conditions, but their figures do not actually reveal its existence However, as shown above, the threshold immigration rate is equal to $\bar{\pi}/\bar{K}$, and it is clear that in their experiments $\bar{\pi}$ was relatively small (probably about 1 per cent. per day) so that the threshold immigration rate is probably also quite small. The lowest rate employed was one mouse in three days, and it is to be noted that in this experiment, the disease remained practically quiescent for a period of 70 to 80 days, apart from a small increase in deaths at one point. During the whole period the size of the population was rising slightly, so that it seems clear that the rate of immigration was just slightly too high to allow the non-specific death rate to keep the population at a constant level, and thus to render a steady state possible An experiment on a larger scale with a proportionally smaller immigration rate would probably reveal the existence of a threshold.

---

* Topley and Wilson "Principles of Bacteriology and Immunity," vol II, p. 767 (1929), and related papers, especially 'J Hyg.,' vol. 24, p. 45 (1925).

† *loc. cit.*, p. 782

We realize that the discussion of epidemics developed in this and in the previous papers is at best only a schematic representation of the invasion of a community by an infective disease, and is far from giving a representation at all adequate or complete The mathematical analysis so far presented may, however, be regarded as first step in the elucidation of the problem from the theoretical point of view. The experimental work of Greenwood and Topley and others gives another complementary method of approach So far the two lines of attack can be brought into relation with each other only very incompletely, but it is to be hoped that as the experimental material becomes more extensive, and the mathematical investigation becomes more comprehensive, a relatively complete understanding of the processes involved in endemic and epidemic invasion may emerge

Summary

(1) The mathematical investigation of the progress of an infectious disease in a community of susceptible individuals has been extended to include the case where members of the community are removed as the result of some general cause of death acting according to constant non-specific death rates, as well as by death from the disease itself. Under the more general conditions here dealt with the main conclusions arrived at in the previous paper remain qualitatively unaltered The limitations which remain are that the susceptibility and the infective power of the individual are supposed to be independent of his age, and further that specific individual immunity does not exist in the sense that the part of the population which escapes infection is assumed to be just as susceptible as the whole population would have been if it had not been infected.

(2) In the general case a unique steady state is found to exist provided that certain relatively simple conditions are satisfied. In the special cases considered a unique steady state in general exists when these conditions continue to be satisfied; but in particular instances, when these conditions are not

satisfied, unique steady states will exist provided that certain other requirements are fulfilled.

(3) Increase of birth rates, in general, increases both the absolute and the relative prevalence of the disease in its steady state. The effect of increase in the non-specific death rates is less simple, but has been worked out at some length.

Decrease in the infectivity of the disease or in the susceptibility of the uninfected results in an increase in the whole population density as well as in an increase in the number of infected. The effect upon the relative incidence of the disease cannot be simply expressed, but it has been worked out in detail in the text. In the absence of immigration, and with the birth rates and also the non-specific death rates equal for virgins and recovered, variation in infectivity or susceptibility will not alter the relative incidence of the disease The total population, however, will increase with decrease of either of these two factors, whilst the number of diseased will also increase proportionately

(4) Two types of threshold values have been encountered. In the first type the quantity in question must initially exceed the threshold value if the event or process is to occur in the population. Two examples of this type have been found, namely, in cases 4 and $2'_a$. In the second type the quantity in question need not initially exceed the threshold value, but may gradually change as the system developes. The event or process to which the threshold refers can only take place when the threshold value has been exceeded. The total population density has a threshold value of this second kind with reference to the existence of steady states.

The Velocity of Sound in Gases in Tubes.

By G. W. C. KAYE, O.B.E., M.A., D.Sc., and G. G. SHERRATT, B.A., Physics Department, National Physical Laboratory, Teddington.

(Communicated by Sir Joseph Petavel, F.R.S.— Received January 27, 1933.)

Introduction.

The main objects of the series of experiments described in the present paper were two-fold: (*a*) to make precision measurements, over a moderate temperature-range, of the velocity of sound-waves in gases contained in resonating tubes; (*b*) to investigate the applicability of the Helmholtz-Kirchhoff expression for the relation between the "free-space" velocity and the "tube" velocity in a gas.

As regards (*b*) different observers in the past have reached diverse conclusions as to the dependence of tube-velocity on the frequency, on the roughness of the wall surface, and on the nature of the material of the tube.

With a view to a fresh examination of the influence of the several factors, six gases of widely differing viscosities and densities were experimented with. Six tubes were used, of various diameters, and of three different materials (glass, copper and carbon) chosen to provide a varied range of such quantities as thermal conductivity and roughness of surface. A comprehensive series of frequencies (500 to 27,000) was available from a valve-oscillator circuit controlled by means of a quartz crystal vibrating piezo-electrically. In order to keep the investigation within reasonable bounds, it was decided to work at two temperatures only, viz., at room temperature and at the steam point.

These experiments form a continuation of previous work at the Laboratory* in which one resonating tube only was used.

Description of Apparatus.

The apparatus consisted essentially of a fixed source of sound mounted immediately opposite the open end of a resonating tube, the acoustic impedance of which could be varied by means of a movable piston. Resonance was detected by means of the reaction on the source. The arrangement has many advantages over the aural method of detecting resonance by means of a listening side-tube and obviates the experimental inconveniences which arise when a movable source is employed. A further advantage lies in the information that is afforded about the form of the sound-waves in the tube

---

* Sherratt and Awbery, 'Proc. Phys. Soc.,' vol. 43, p. 242 (1931).

A section of the main portion of the apparatus is shown in fig. 1. The resonating tube was mounted by means of brass rings inside a longer brass tube, one end of which carried a gastight cylindrical enclosure within which was mounted the source of sound The other end of the brass tube was fitted with a hollow cylindrical "trap" provided with two packing glands which permitted the traversing of a steel "piston" rod. This "trap" served the purpose of reducing the possibility of air leakage past the two glands into the main body of the apparatus

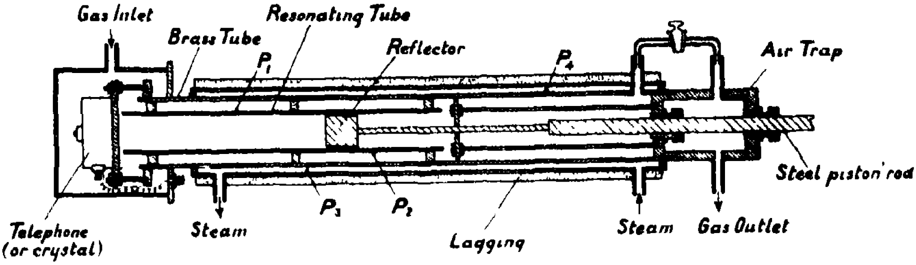


Fig. 1 Sectional diagram of apparatus

The inner end of the steel rod carried a cylindrical reflector having its face accurately plane and perpendicular to its axis, while to the outer end was attached a vernier moving over a steel centimetre scale. The accuracy of this scale was verified in the Metrology Department of the Laboratory as being abundantly adequate for the present purpose. The range of movement of the reflector, amounting to 60 cm., was restricted to the central portion of the resonating tube so as to ensure that measurements taken at the steam point were well within the uniform temperature region. The apparatus was enclosed for almost the whole of its length within a lagged steam-jacket and was mounted on an optical V-bench for ease of assembly and adjustment. The long brass tube enveloping, as it did, a considerable length of the piston rod served to minimize undesirable heat loss by conduction along the rod when observations were being taken at the steam point.

Temperature measurements of the gas under experiment were made by means of two calibrated copper-constantan thermocouples P_1 and P_2 soldered on to brass sleeves cemented in good thermal contact with the outside of the resonating tube, except for the copper tubes when the couples were soldered on to the tubes themselves. The sleeves were situated one towards each end of the range traversed by the reflector. Two similar thermocouples P_3 and P_4 were also mounted inside the steam-jacket itself.

The design of the apparatus permitted the ready interchange of the various resonating tubes without affecting the alignment of the essential parts. Two experimental precautions may be mentioned here It was pointed out in the discussion which followed the paper by Sherratt and Awbery (*loc cit.*) and also in a recent paper by Henry* that the gas behind the piston, as well as that in front, may form a vibratory system. As a result small errors may be introduced if the piston is not a good fit in the tube. Another possible source of error lies in the longitudinal vibration of the piston rod To prevent appreciable errors from these sources the reflecting cylinders were made especially heavy and with no more clearance than was necessary to allow them to be traversed in the tubes. The several cylinders used were between 3 and 5 cm. long, and the weights ranged from 8 to 245 grams

Sound Production.

Sound waves were produced by either of two methods —

(i) By means of a telephone supplied with alternating current from an oscillatory valve-circuit, controlled in frequency by a quartz crystal vibrating piezo-electrically ,

(ii) By means of the same quartz crystal operating directly on the resonating column.

The quartz crystal used for the purpose was cut in the form of a square sectioned rod (10 × 1·5 × 1·5 cm.) the optic axis being along the length of the rod, and two of the longitudinal faces being at right angles to one of the electric axes.

Fig. 2 shows the circuit diagram for the telephone method. It consists essentially of four separate circuits, marked off by dotted lines in the diagram. Circuit I is a two-valve master-oscillator and is an adaptation of the Abraham-Bloch multi-vibrator circuit controlled in frequency by the quartz crystal vibrating in a flexural mode as described by Harrison† and others. For the purpose the crystal rod was clamped between two pairs of parallel metal strip electrodes mounted along the longitudinal edges. The practical advantage of this two-valve over the more usual one-valve circuit lies in the fact that

* 'Proc. Phys Soc.,' vol 43, p. 340 (1931).

† 'Proc. Inst Radio Engrs.,' vol. 15, p. 1040 (1927).

much smaller inductances and capacities are required when comparatively low frequencies are being dealt with.

Circuit II is a multi-vibrator circuit similar to circuit I and serves to provide an alternating current which is rich in harmonics and is controlled in frequency by circuit I. If n is the natural frequency of oscillation of the crystal, the frequencies obtainable in circuit II by variation of the condensers C_1 and C_2 (fig. 2) are $n/2$, $n/3$, $n/4$ and so on down to $n/15$. It was found in practice that the condenser settings required to tune the circuit to any one of the sub-harmonics were not critical.

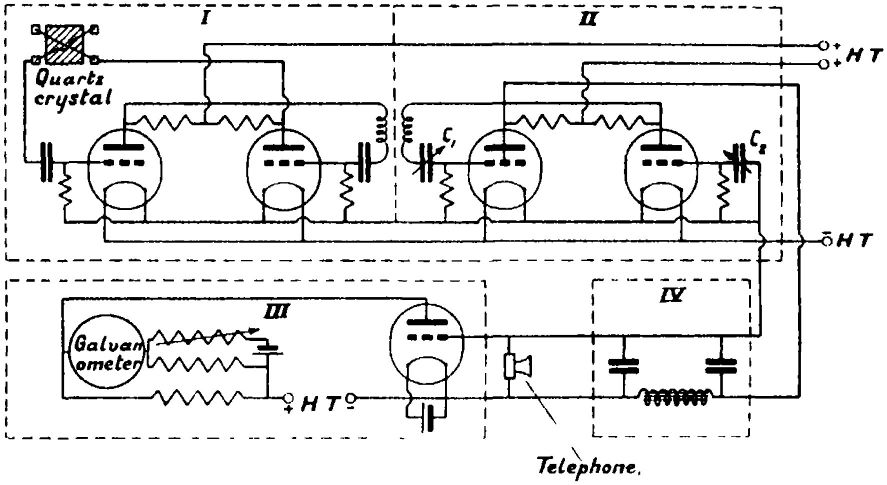


FIG 2 —Circuit diagram.

I - Master-oscillator circuit
II—Multi-vibrator circuit.
III—Rectifying circuit.
IV—Low-pass filter.

Circuits I and II together thus constitute a means of producing a series of accurately known frequencies. These are fed into the telephone, the e.m.f. across the telephone terminals being rectified by circuit III which is a single-valve circuit employing anode-bend rectification. The main part of the plate current was balanced out by means of a potentiometer, the remainder being passed through a sensitive galvanometer. Changes in the rectified plate-current are a measure of variations in the e.m.f. across the telephone, which e.m.f. will in turn depend on the amplitude and frequency of the oscillations in circuit II, and on the mechanical impedance of the telephone diaphragm as well as on the acoustical impedance of the gas enclosures in front of and behind the telephone diaphragm. During a particular experiment the frequency was kept constant, the only quantity varied being the length and so the impedance of the gas column. If these variations in the impedance form an appreciable fraction of the gross impedance of the vibrating system they will be rendered evident in the plate current in circuit III. In effect this means that as the reflector is moved along the resonating tube, the galvanometer deflection will pass through a series of maxima and minima, the distance between positions corresponding to successive maxima or minima being equal to half the wavelength of the sound in the tube at the particular frequency used.

Fig 3 is a typical curve showing the galvanometer current plotted against the position of the movable reflector. The greater sharpness of the maxima compared with that of the minima is because the impedance of the vibrating system changes most rapidly with the reflector setting when resonance occurs.

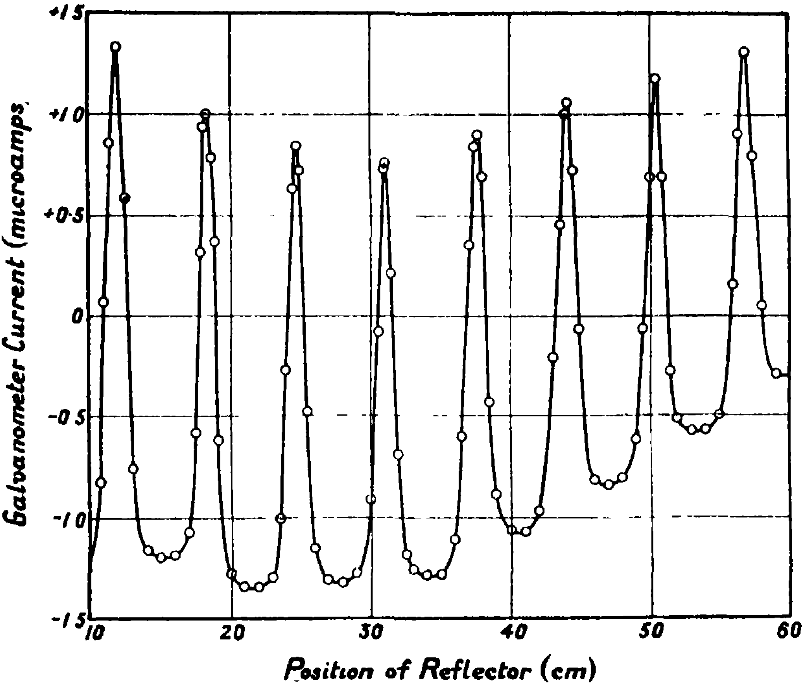


FIG 3.—Air in glass tube A II Frequency 2636 cycles/sec. Temperature 16 7° C

For frequencies below $n/8$, it was found necessary to include a simple low-pass filter (circuit IV) to ensure an approximately sinusoidal wave form which otherwise was vitiated by the presence of the first harmonic. Frequency $n/2$ could not be used on account of the lack of sensitivity of the telephone at such frequencies.

The second method of sound production was to use the quartz crystal itself as a source of sound. The crystal was mounted in the same manner as the telephone, opposite the end of the resonating tube. Two modes of vibration were employed: the flexural mode described above and the "transverse" mode. In the flexural mode the quartz rod was mounted "broadside on," while in the case of the "transverse" mode the rod was mounted "end on" opposite the end of the resonating tube.

For the flexural mode of vibration, circuit I of fig. 2 was used as before. Circuits II and III were dispensed with, the galvanometer and its associated potentiometer being included directly in circuit I to indicate variations in the plate current. Curves similar in shape to fig. 3 were obtained, except that the maxima were now even sharper.

In the "transverse" mode of vibration, the crystal was clamped between metal electrodes placed on the two faces perpendicular to the electric axis. The motion consists of longitudinal vibrations along the length of the quartz bar, with a node at the mid-point. For the purpose a simple one-valved circuit, such as was first used by Pierce, was employed, the crystal being connected between the grid and filament of the valve with a tuned inductance in the plate circuit.

The natural frequencies of the two modes of vibration of the quartz crystal were determined with high precision in the Electrical Department of the Laboratory. The natural frequency in its first flexural mode was 7907·95 cycles per second, and in the "transverse" mode 27,422 cycles per second. It was verified during the determination that the frequency of vibration of the quartz was unaffected by changes in the acoustic load. No special precautions are required to keep the frequency constant, except that the gaps between the crystal and its electrodes must be maintained constant, this being ensured by means of ebonite spacers which also serve to clamp the crystal. The temperature coefficient of frequency was very small (a few parts in a million per degree Centigrade).

Resonating Tubes and Gases Used.

Six resonating tubes were used in the present experiments. Particulars of the material, bore and wall thickness are given in Table I.

Tubes A I and A II were of Wood's soda glass. B I and B II were solid drawn copper tubes. Carbon tube C I was included because it represents the type of tube which is to be used in forthcoming experiments on the measurement of the velocity of sound in various gases at high temperatures. Neither C I nor C II was perfectly straight.

Table I.

Material:	Glass.		Copper.		Carbon.	
Tube:	A I.	A II.	B I.	B II.	C I	C II.
	cm.	cm.	cm.	cm.	cm.	cm.
Mean internal diameter	$2{\cdot}89_5$	$1{\cdot}04_7$	$2{\cdot}21_4$	$0{\cdot}92_4$	$1{\cdot}29_2$	$0{\cdot}88_5$
Mean wall thickness	0·19	0·10	0·16	0·18	0·61	0·45

The internal diameters of the tubes were measured by the method devised by Rolt, Wilmotte and McPetrie.* It consists in effect in measuring the distance between the centres of two accurate steel balls of known diameter when pressed lightly in contact with each other and with the walls of the tube. The method permits the accurate determination of the bore of the tube at any point along its length. Measurements were made at intervals of 1 inch over the range traversed by the movable reflector. The means of these are given in Table I. The maximum deviations from the mean diameter in the order of the results given in Table I were :—

0·01, 0·02, 0·005, 0·003, 0·02 and 0·03 cm.

The glass and carbon tubes were slightly oval in section, the copper tubes being much more uniform in diameter.

The gases used were dry air free from carbon dioxide, hydrogen, carbon dioxide, sulphur dioxide, ammonia and ethyl chloride. With the exception of the air they were all contained in gas cylinders. The air was passed through soda lime, caustic potash, calcium chloride and phosphorus pentoxide to remove any carbon dioxide and moisture present. The ethyl chloride was supplied as containing no impurity, and the ammonia as containing less than 0·005 per cent. impurity. The remaining gases were analysed by the Chemistry Division of the Laboratory. The sulphur dioxide was found to contain less than 0·01 per cent. impurity, while the analyses of the carbon dioxide and hydrogen were as follows :—

---

* 'J. Sci. Instr.,' vol. 6, p. 379 (1929).

Carbon dioxide.

	per cent.
CO_2	99·08
N_2	0·85
H_2	0·04
O_2	0·02
CO	0·01

Hydrogen.

	per cent.
H_2	99·34
N_2	0·58
O_2	0·07
CO	0·01

Experimental Procedure

The apparatus was first exhausted by means of a vacuum pump connected to the gàs outlet tube, the exterior packing-gland round the piston being temporarily waxed over to prevent leaks through the packing material. To assist in drying out the apparatus steam was passed through the steam-jacket. The pump was then stopped and the pressure inside the apparatus raised to atmospheric by admitting a supply of the required gas. The procedure was repeated several times and finally a slow stream of gas was passed through the apparatus continuously, thus maintaining the pressure slightly above atmospheric, and so preventing the entrance of air via the exterior packing gland.

By the use of an exploring shielded thermocouple inside the resonating tube, the effect of the gas flow on the temperature distribution along the tube was determined, and also the difference between the actual temperature of the gas and that measured by the thermocouples P_1 and P_2 attached to the outside of the tube. At room temperature no difference on this score was, of course, to be anticipated. At the steam point, also, it was verified that the temperature of the gas within the range of motion of the reflector was satisfactorily uniform, there being a small progressive drop in temperature towards the telephone end amounting to about 0·2° C. except for the copper tubes and hydrogen gas, when the temperature drop was about 0·5° C. The readings of the thermocouples P_1 and P_2 differed by less than these amounts from each other and from the readings of the exploring thermocouple inside the tube. Thus the mean temperature of the gas at the steam point could be assumed to be known to 0·2° C. even under the least favourable conditions.

The apparatus having been filled with gas and the temperature being everywhere steady, the valve circuits were excited and the telephone or crystal was set in vibration. The galvanometer reading was adjusted to a convenient amount by means of the potentiometer, time being allowed for the battery voltages to become steady. The piston was then slowly traversed and the resulting galvanometer deflections noted. The procedure found best for determining the positions of the maxima was to approach each of them first from one side and then from the other, as exemplified in the two adjacent columns of readings for each node enumerated in Table II. The true position of a maximum was taken to be the mean of the two readings so obtained. These differed by about 0·5 mm. for the shortest wave-length (about 3 cm.) and by not more than 5 mm. at the longest wave-length (about 80 cm.) used. Several pairs of such readings were obtained for each maximum. No attempt was made to determine the exact positions of the comparatively broad minima.

The " end correction " was avoided by the usual method of differences, for which purpose attention was normally confined to a relatively few nodes situated towards the two ends of the range traversed by the reflector. For each new condition of temperature or frequency, however, the positions of a number of nodes distributed over the complete range were first determined. This was done to verify that the wave-length did not vary with distance as might happen if the sound-intensity were too large. Actually it was found that the nodes were equidistant within the limit of error over the whole range for all the gases and for all the frequencies which could be employed under the particular conditions.

Two practical points may be noted here. With hydrogen, the velocity of sound is so high that even the highest available frequencies (7908 and 27,422 cycles per second) only yielded barely sufficient maxima for purposes of wave-length measurement. But the use of these very high frequencies proved to be restricted for another reason. With any other gas than hydrogen it was not possible, with any of the tubes, definitely to locate the maxima when using the highest frequency (27,422). A similar state of things obtained with the two widest tubes (*i.e.*, A I and B I), and frequency 7908 except when the tubes were filled with hydrogen or ammonia. The " nodal " curves no longer displayed single sharp maxima, as in fig. 3, but rather a number of secondary maxima. The presence of such " satellites " with very high frequencies was attributed by Pielemeier* to successive reflections at the quartz crystal and the reflector, the velocity of propagation of the sound decreasing with the amplitude. That this cannot be the explanation in the present experiments is indicated by the fact that the satellites disappear if the diameter is reduced, or if the wave-length is increased (either by lowering the frequency or by changing the gas to one in which the velocity is higher).

We are thus led to think that the effect owes its origin to the fact, first indicated by Rayleigh,* that a non-plane wave of sound entering a tube does not become plane as it travels along the tube, unless the wave-length exceeds 3·4 times the radius. If the ratio of wave-length to radius is smaller than this, radial vibrations of the gas column are set up to the detriment of precise nodal location.

---

* 'Phys. Rev.,' vol. 34, p. 1184 (1929).

In all experiments the observed velocities at room temperature and the steam point were reduced to those at 18° C. and 100° C. respectively, by assuming the velocity to be proportional to the square root of the absolute temperature and neglecting any change in γ, the ratio of the specific heats, over the small temperature ranges involved. The readings of the four thermocouples were taken both at the beginning and end of each experiment.

Typical Experiment.

The results of a typical experiment are reproduced below in full.—

Carbon dioxide in Carbon Tube C I.

$$\text{Frequency} = \frac{n}{4} = \frac{7908}{4} = 1977 \text{ cycles per second.}$$

Thermocouple.	Microvolts.	
	Initial	Final.
P_1	637	637
P_2	639	639
P_3	641	642
P_4	642	642

Mean temperature (from P_1 and P_2) = 17·0° C.

Table II.

Position of 1st node.		Position of 2nd node.		Position of 3rd node.	
11·90 cm.	12·04 cm.	18·59 cm.	18·76 cm.	25·24 cm.	25·40 cm.
11·92 "	12·03 "	18·60 "	18·80 "	25·24 "	25·40 "
11·90 "	12·05 "	18·57 "	18·77 "	25·25 "	25·41 "
11·91 "	12·05 "	18·58 "	18·76 "	25·26 "	25·40 "
11·92 "	12·03 "	18·58 "	18·76 "	25·23 "	25·40 "
11·92 "	12·04 "	—	—	—	—
11·91 cm.	12·04 cm.	18·58 cm.	18·77 cm.	25·24 cm.	25·40 cm.
Mean 11·975 cm.		Mean 18·675 cm.		Mean 25·32 cm.	

---

* "Theory of Sound," vol. 2, p. 161 (1894).

Table II.—(continued).

Position of 7th node.		Position of 8th node.		Position of 9th node.	
52·11 cm.	52·30 cm	58·78 cm.	58·95 cm.	65·49 cm.	65·66 cm.
52·12 ,,	52·34 ,,	58·82 ,,	58·97 ,,	65·50 ,,	65·67 ,,
52·12 ,,	52·30 ,,	58·79 ,,	58·95 ,,	65·49 ,,	65·65 ,,
52·12 ,,	52·29 ,,	58·79 ,,	58·95 ,,	65·49 ,,	65·65 ,,
52·12 ,,	52·30 ,,	—	—	65·49 ,,	65·64 ,,
52·12 cm.	52·31 cm.	58·79 cm.	58·95 cm.	65·49 cm.	65·65 cm.
Mean 52·215 cm		Mean 58·87 cm.		Mean 65·57 cm.	

$$52{\cdot}215 - 11{\cdot}975 = 40{\cdot}24 \text{ cm.}$$

$$58{\cdot}87 - 18{\cdot}675 = 40{\cdot}195 \text{ cm.}$$

$$65{\cdot}57 - 25{\cdot}32 = 40{\cdot}25 \text{ cm.}$$

$$\text{Mean} = 40{\cdot}228 \text{ cm.}$$

$$\text{Half wave-length at } 17{\cdot}0^\circ \text{ C.} = \frac{40{\cdot}228}{6} = 6{\cdot}705 \text{ cm.}$$

$$\text{Correcting to } 18^\circ \text{ C., half wave-length} = 6{\cdot}705 \times \left(\frac{291}{290}\right)^{\frac{1}{2}}.$$

$$= 6{\cdot}716 \text{ cm.}$$

Tube-velocity in CO_2 at 18° C. for a frequency of 1977 cycles per second $= 6{\cdot}716 \times 2 \times 1977$ cm per second $= 265{\cdot}6$ metres per second.

Results.

The whole of the tube-velocity results are given below in Tables IIIA and IIIB, each value being the mean of several determinations derived as exemplified above. The deviations from the mean are of the order of 1 part in 1000 for the lowest frequencies and 1 part in 2000 for the highest.

Discussion of Results.

The theoretical equation obtained by Helmholtz and modified by Kirchhoff states that if V_0 represents the velocity of sound of frequency n in free space and V that in the same gas in a tube of radius r, then

$$V = V_0 \left\{1 - \frac{c}{2r\,(\pi n)^{\frac{1}{2}}}\right\}, \tag{1}$$

Table IIIA.

Tube velocities (V) (metres/sec.)

(n = 7908 cycles/sec.)

Tube.	Frequency	Air.		CO_2.		SO_2		NH_3.		C_2H_5Cl.	
		18° C	100° C	18° C	100° C	18° C	100° C.	18° C	100° C	18° C	100° C.
Glass	n	—	—	—	—	—	—	427·7	481·7	—	—
	$n/3$	341·8	386·4	266·4	297·7	216·0	244·1	427·5	480·9	203·7	230·4
	$n/4$	341·6	386·3	266·3	297·7	215·8	243·8	427·6	481·2	203·5	230·1
	$n/5$	341·7	386·0	266·4	297·7	215·8	243·8	427·4	480·5	203·5	230·2
	$n/6$	—	386·2	266·2	297·6	216·0	243·8	427·0	480·8	203·5	—
	$n/7$	341·4	—	266·2	297·6	215·9	243·6	427·3	480·2	—	—
A I	$n/8$	—	385·7	—	297·4	216·0	243·7	—	480·2	—	230·1
	$n/9$	341·4	—	—	—	—	—	—	—	203·3	—
	$n/10$	341·2	385·7	266·2	297·4	215·7	243·5	427·1	479·9	203·0	230·1
	$n/11$	341·0	—	—	—	215·7	243·7	—	—	203·3	230·1
	$n/12$	—	385·4	266·0	297·3	215·8	243·4	—	479·6	202·9	—
	$n/15$	340·8	—	—	—	—	—	426·9	480·1	—	229·7
	n	341·3	385·9	266·2	297·5	215·8	243·7	427·2	480·0	203·3	229·9
	$n/3$	340·4	384·8	265·7	296·9	215·6	243·3	426·1	479·3	203·4	229·9
	$n/4$	340·5	384·3	265·7	296·9	215·4	243·1	426·1	478·2	203·1	229·7
	$n/5$	340·1	384·0	265·6	296·6	215·5	243·3	425·9	478·3	203·1	—
	$n/6$	339·8	384·0	265·6	296·3	215·2	242·9	425·5	478·0	—	229·3
	$n/7$	—	383·5	265·5	—	—	243·1	—	477·5	202·5	229·4
	$n/8$	339·4	383·3	—	296·1	—	—	424·8	—	202·6	—
A II	$n/9$	339·0	—	265·2	—	215·2	—	—	—	202·4	229·2
	$n/10$	339·1	382·8	265·2	296·0	215·2	242·7	425·0	477·3	202·5	229·2
	$n/11$	—	—	—	—	—	—	—	—	—	—
	$n/12$	338·7	—	265·0	295·7	214·8	—	424·6	476·3	—	228·7
	$n/15$	—	—	—	—	214·9	—	—	—	201·8	—
Copper	n	341·9	386·7	—	—	—	—	427·6	481·1	—	—
	$n/3$	341·5	386·0	266·4	297·6	215·9	243·9	427·4	480·6	203·5	230·1
	$n/4$	341·2	386·0	266·1	297·5	215·8	243·7	427·0	480·6	203·5	230·0
	$n/5$	341·3	385·4	266·2	297·5	215·7	243·7	427·1	480·1	203·2	230·0
	$n/6$	341·2	385·5	266·1	297·3	215·7	243·6	427·0	—	203·2	229·9
	$n/7$	—	385·6	—	297·4	—	243·7	426·7	—	—	—
B I	$n/8$	341·0	—	266·0	—	215·8	243·5	—	479·4	203·2	230·0
	$n/9$	—	—	—	297·3	—	—	426·4	—	202·9	229·8
	$n/10$	340·8	385·1	266·0	297·2	215·6	243·6	426·4	479·7	202·8	229·8
	$n/11$	—	—	—	—	—	—	—	—	—	—
	$n/12$	340·6	384·7	265·8	—	215·7	243·2	426·5	479·7	202·6	229·7
	$n/15$	340·3	—	—	296·9	—	243·3	—	—	—	229·4
	n	341·2	385·6	266·1	297·4	215·7	243·6	426·9	480·2	203·2	229·6
	$n/3$	340·4	384·3	265·7	296·9	215·5	243·2	426·0	479·0	203·2	229·5
	$n/4$	340·0	384·0	265·5	296·5	215·4	243·2	425·5	478·5	203·0	229·5
	$n/5$	339·9	383·6	265·4	296·3	215·4	242·8	425·4	478·0	202·9	—
B II	$n/6$	339·4	383·4	265·1	—	215·2	242·9	425·2	—	—	229·1
	$n/7$	—	—	265·1	296·0	—	242·8	424·9	—	202·5	229·2
	$n/8$	—	382·5	265·1	295·7	215·2	242·5	424·4	—	202·3	229·1
	$n/9$	339·0	—	264·8	—	—	—	—	—	—	—
	$n/10$	338·6	382·3	264·9	295·4	215·0	242·3	424·7	476·4	202·0	228·8
	$n/11$	—	—	—	—	—	242·1	—	—	201·9	228·9
	$n/12$	338·2	381·8	264·5	295·5	214·7	242·5	—	—	201·7	228·6

Table IIIA—(continued).

Tube	Frequency	Air		CO_2.		SO_2.		NH_3.		C_2H_5Cl.	
		18°C.	100° C	18° C	100° C	18° C	100° C	18° C	100° C	18° C.	100° C
Carbon	*n*	341·2	385·5	266·2	297·4	215·8	243·7	426·9	480·2	—	—
	n/3	340·4	384·5	265·6	296·7	215·4	243·2	425·7	478·5	-	—
	n/4	339·8	384·1	265·6	296·8	215·5	243·1	425·8	478·6	—	—
	n/5	339·8	383·4	265·2	296·3	215·4	242·8	425·4	478·0		—
	n/6	-	383·5	265·4	296·2	215·0	243·1	425·1	477·3	—	—
	n/7	—	—	265·1	295·6	215·1	242·5	—	476·5	—	—
	n/8	339·1	382·5	265·0	295·6	—	242·5	424·9	476·5	—	—
C I	*n*/9	338·7	382·5	—	—	—	—	—	—	—	—
	n/10	338·5	382·0	264·9	295·5	215·1	242·5	423·7	476·4	—	—
	n/11	338·5	—	—	—	—	—	—	476·1	—	—
	n/12	337·9	381·3	—	295·0	215·0	241·9	424·3	474·9	—	—
	n/15	—	380·5	264·1	294·9	214·4	—	422·5	—	-	—
	n	340·8	384·9	265·9	297·0	215·6	243·4	426·3	479·2	—	—
	n/3	339·2	383·0	265·2	296·1	215·1	242·7	425·0	477·3	—	—
	n/4	339·0	382·3	265·1	295·8	214·9	-	424·3	476·3	—	—
	n/5	338·5	382·1	264·6	295·2	215·0	242·5	424·1	476·5	—	—
	n/6	337·9	—	264·5	295·3	214·6	242·0	423·1	474·9	—	—
C II	*n*/7	337·6	381·0	264·2	—	214·6	242·0	422·8	—	—	—
	n/8	—	380·4	—	294·9	214·8	—	423·3	—	—	—
	n/9	337·2	380·1	264·3	—	214·2	241·3	423·0	—	—	—
	n/10	336·3	379·5	264·0	294·1	214·3	241·5	422·6	473·4	—	—
	n/11	—	379·5	—	—	—	—	421·6	—	-	—
	n/12	336·3	379·3	263·4	—	214·3	-	421·5	—	—	—

Table IIIB.

Tube velocities (V) (metres/sec.)

Tube	Frequency (cycles/sec.).	Hydrogen	
		18° C.	100° C
Glass—			
A I	27422 7908	1296 1292	1457 1455
A II	27422 7908	1293 1286	1456 1445
Copper—			
B I	27422 7908	1294 1291	1458 1453
B II	27422 7908	1291 1284	1454 1445
Carbon—			
C I	27422 7908	1293 1285	1456 1442
C II	27422 7098	1290 1280	1454 1437

where c is the Kirchhoff constant given by

$$c = \mu^{\frac{1}{2}} + \left(\frac{\nu}{\gamma}\right)^{\frac{1}{2}} (\gamma - 1), \tag{2}$$

where

μ denotes the kinematic viscosity,
γ is the ratio of the specific heats, and
ν is the thermal diffusivity of the gas.

Now, although certain features of the Helmholtz-Kirchhoff equation have been subjected to criticism in the past, practically all experimenters are in agreement as to the linear variation of V with the reciprocal of r, provided that the comparison is restricted to tubes which have as nearly as possible the same kind of inner surface. We may thus assume that for a given frequency, gas and temperature, two tubes of the same material and different radii (r_1, r_2) will be respectively associated with tube velocities (V_1, V_2) which are related to the free-space velocity by the expression

$$V_0 = (V_1 r_1 - V_2 r_2)/(r_1 - r_2). \tag{3}$$

The procedure adopted in applying this equation to the present results was to take one particular gas and to plot the frequency against the tube velocity at 18° C. for say each of the two glass tubes. The resulting smooth curves enable the respective tube velocity corresponding to each of a series of selected frequencies to be read off. We thus ascertain V_1 and V_2 for each frequency and, as r_1 and r_2 are known, we can calculate V_0. Similarly by treating in turn the results at 18° C. for the two copper and the two carbon tubes, two other values of V_0 are obtained for each of the selected frequencies. If the assumed velocity-radius relation is correct, the various values of V_0 should show no appreciable variation with either frequency* or tube material. The same process is now repeated for the results at 100° C., and finally the whole procedure is carried out for each of the other gases, except that for hydrogen, smoothed velocities cannot be derived in view of the fact that only two frequencies could be used experimentally.

In Tables IVA and IVB are given the values of V_0 obtained in this manner for a variety of frequencies. The probable accuracy is of the order of 1 in 1000 except for hydrogen where it may be a little less.

It will be seen that with none of the gases is there any significant variation of V_0 with frequency or tube material and so the truth of the linear variation

---

* Apart, of course, from any possible variation of free-space velocity with frequency.

Table IVa.

Calculated Free-space Velocities (V_0) (metres/sec.)

(n = 7908 cycles/sec.)

Tube.	Frequency	Air.		CO_2		SO_2.		NH_3.		C_2H_5Cl.	
		18° C.	100° C	18° C.	100° C	18° C.	100° C	18° C	100° C.	18° C.	100° C.
Glass	1 0 n	—	—	—	—	—	—	428 2	481 9	—	—
	0 4 n	342 4	387 5	266 7	298 2	216 2	244·2	428·3	482·1	204 0	230 7
	0 3 n	342·4	387 4	266 7	298·3	216 3	244 3	428 3	482 2	203 9	230·7
	0 2 n	342 4	387 3	266 7	298·1	216·2	244 2	428·2	482·0	203·8	230·5
	0·1 n	342 3	387·0	266·6	298 3	216 2	244 1	428·3	481 8	203·7	230·6
Copper	1 0 n	342·3	387 3	—	—	—	—	428·2	481·9	—	—
	0 4 n	342 4	387 2	266 6	298·1	216 1	244 2	428 2	481·8	203 9	230 5
	0 3 n	342 3	387 2	266 7	298 2	216 1	244 2	428 2	481·8	203 9	230·6
	0·2 n	342 2	387 2	266 7	298 2	216·2	244·2	428 2	481 7	203 7	230 5
	0·1 n	342·3	386 9	266·7	298 2	216 2	244·2	427 9	481 6	203·5	230 6
Carbon	1·0 n	342 4	387 3	266 7	298 2	216 2	244 1	428 2	481 8	—	—
	0·4 n	342 3	387 7	266 7	298 2	216 4	244 3	428 1	482 0	—	
	0 3 n	342 3	387 5	266 7	298 3	216·2	244 3	428 1	482 0	—	-
	0 2 n	342 5	387 4	267 0	298 3	216 0	244 2	428 2	482 3	—	-
	0·1 n	342·5	386 8	266 7	297 6	215 9	244 0	428 5	481 5	—	—
Mean V_0		342 4	387 3	266 7	298 2	216 2	244 2	428 2	481 9	203 8	230·6
Corrected mean V_0		342 4	387 3	265 8	297 2	216 2	244 2	428·2	481 9	203 8	230·6

Table IVb.

Calculated Free-Space Velocities (V_0) (metres/sec)

Tube.	Frequency (cycles/sec).	Hydrogen.	
		18° C.	100° C.
Glass	27422 7908	1297·7 1295 5	1457·6 1460 7
Copper	27422 7908	1296 1 1295·9	1460 8 1458 5
Carbon	27422 7908	1299 7 1296 2	1460 5 1453 2
Mean V_0		1297	1458·5
Corrected mean V_0		1301	1463

of V with $1/r$ is amply confirmed. The last line in Table IV "Corrected mean V_0" was obtained after applying a correction for the effect of the impurities known to be present in the carbon dioxide and hydrogen. This correction was evaluated by applying the basic formula $V_0 = \sqrt{\gamma P/\rho}$ in turn to the impure gas and the ideally pure gas. The density ρ at pressure P of the impure gas was calculated from those of its constituents, assuming the truth of Dalton's law, whilst γ the ratio of specific heats for a mixture was taken to be related to the values for the components, by the formula of Richarz.*

Attention may be directed to the result for carbon dioxide at 100° C. in view of its bearing on the molecular structure of that gas. Two models have been proposed for the CO_2 molecule, a "straight" model with the three atoms in a straight line, and a "bent" model in which they are placed at the corners of a triangle The specific heats corresponding to the two forms were calculated by Eucken,† whence it appeared that while at room temperatures the two values agreed too closely for experiment to differentiate between them, the results at low temperatures favoured the straight model. At high temperatures the experimental information was conflicting, calorimetric measurements supporting the straight model, while existing sound-velocity determinations afforded specific heats more in agreement with the bent model. Later Rawlins‡ who reviewed the data provided by infra-red absorption spectra and other physical properties summed up in favour of the straight model, and recently Ibbs and Wakeman§ from their measurements of the viscosity of carbon dioxide and of the thermal diffusion in mixtures of gases containing carbon dioxide have inferred that the CO_2 molecule is straight at any rate at temperatures up to 145° C. The present experiments on the sound-velocity at 100° C. are the first, so far as we are aware, to give definite support to the straight model.

Having now derived for the different gases, the values of V and V_0 for each frequency and tube material, the next step is to calculate the constant c in the Helmholtz-Kirchhoff equation. This is done in Table V. Naturally no great precision can be anticipated in the value of c, depending as it does, on the difference of two nearly equal velocities.

We have also included in Table V the "theoretical" values of c derived from

---

* 'Ann. Physik,' vol. 19, p. 639 (1906).

† 'Z. Physik,' vol. 37, p. 714 (1926).

‡ 'Trans. Faraday Soc.,' vol. 25, p. 925 (1929).

§ 'Proc. Roy. Soc.,' A, vol. 134, p. 613 (1932).

Kirchhoff's expression (2) given on p. 136. The values used for the viscosity, density, thermal conductivity and specific heat are those given in the International Critical Tables. No data for the viscosity of ethyl chloride could be traced for the temperature range required. Hydrogen is also omitted from the table in view of the impracticability of deriving smoothed values for the sound velocities.

Table V.

Experimental Values of the Kirchhoff Constant (c).

($n = 7908$ cycles/sec.)

Tube	Frequency.	Air.		CO_2		SO_2		NH_3		C_2H_5Cl	
		18° C	100° C	18° C	100° C.	18° C	100° C.	18° C	100° C.	18° C.	100° C.
Glass	1·0 n	0·49	0·67	0·31	0·35	0·31	0·29	0·40	0·57	0·62	0·65
	0·4 n	0·51	0·52	0·32	0·39	0·27	0·29	0·40	0·48	0·71	0·68
	0·3 n	0·51	0·58	0·28	0·42	0·23	0·31	0·41	0·47	0·62	0·73
	0·2 n	0·48	0·63	0·31	0·41	0·24	0·33	0·43	0·55	0·59	0·65
	0·1 n	0·55	0·63	0·32	0·39	0·27	0·38	0·40	0·57	0·65	0·74
	Mean	0·51	0·61	0·31	0·39	0·26	0·32	0·41	0·53	0·64	0·69
Copper	1·0 n	0·47	0·67	0·30	0·44	0·33	0·39	0·41	0·60	0·71	0·89
	0·4 n	0·51	0·65	0·33	0·41	0·30	0·31	0·44	0·52	0·65	0·83
	0·3 n	0·53	0·64	0·32	0·42	0·31	0·31	0·44	0·51	0·61	0·78
	0·2 n	0·54	0·64	0·32	0·39	0·29	0·32	0·43	0·55	0·70	0·75
	0·1 n	0·51	0·65	0·33	0·39	0·28	0·34	0·45	0·55	0·76	0·74
	Mean	0·51	0·65	0·32	0·41	0·30	0·33	0·43	0·55	0·69	0·80
Carbon	1·0 n	0·75	0·93	0·46	0·52	0·34	0·51	0·65	0·85	—	—
	0·4 n	0·72	0·84	0·46	0·57	0·36	0·48	0·67	0·78	—	
	0·3 n	0·75	0·87	0·46	0·57	0·36	0·46	0·63	0·77	—	—
	0·2 n	0·75	0·83	0·45	0·58	0·38	0·47	0·62	0·78	—	—
	0·1 n	0·74	0·90	0·49	0·63	0·40	0·53	0·64	0·81	—	—
	Mean	0·74	0·87	0·46	0·57	0·37	0·49	0·64	0·80	—	—
Theoretical value of c		0·56	0·69	0·37	0·47	0·27	0·34	0·48	0·62	—	—

Referring to Table V it will be seen that for the same gas at the same temperature and the same tube material the Kirchhoff constant shows no significant variation with frequency. This means that the tube reduction in sound velocity obeys the $n^{-\frac{1}{2}}$ law as required by the Helmholtz-Kirchhoff formula. With reference to the influence of the tube material, we may note that for the glass and copper tubes the observed values of c are on the average some 10 per cent. smaller than the theoretical values, which in the circumstances may be regarded as reasonably fair agreement. Henry (*loc. cit.*) in considering the possible sources of error in the resonant tube method concluded that the effect in every case would be to decrease the observed value of the Kirchhoff constant; and it may be that the several factors in question are operative in practice to a greater or less extent.

The Helmholtz-Kirchhoff formula would thus appear to be approximately valid for the glass and copper tubes which, though possessing very different physical properties in general, all have smooth surfaces. For the carbon tubes, however, the observed values of c for all the gases are, on the average, approximately 30 per cent. larger than the theoretical values. Support is thus given to the conclusions of previous observers who found that the speed of sound waves in gases in tubes with rough surfaces is less than in tubes of the same diameter with smooth surfaces.

A possible explanation would seem to be that slight irregularities, such as are present in a rough-surface tube, may produce disturbances in the motion of the gas close to the walls. As E. G. Richardson pointed out in the discussion of the paper by Henry (*loc. cit.*) the amplitude of vibration may actually be greater at a small distance from the walls than at the centre of the tube, so that quite small surface irregularities may conceivably give rise to an annular region of irregular disturbance, which would have the effect of virtually decreasing the diameter of the tube. To explain the results obtained in the present experiments with carbon tubes, the thickness of this annular layer would have to be of the order of 1 or 2 mm. We may compare the recent photographs by Andrade* of the tracks of small particles in a Kundt tube. These show that under certain conditions the whole interior of the tube can be given up to an irregular motion, though presumably larger amplitudes than were used in the present experiments would be required to produce such an effect. In any event it is always possible that the phenomena observed by Andrade may be caused, in large part, by the presence of the dust particles themselves.

With a view to throwing light on the matter the particle velocity in air was measured in the centre of the glass tube A I by introducing a Rayleigh disc, the determination being made under conditions of resonance. No observations were possible for the highest frequency (27,422 cycles per second), but for the frequency $n = 7908$ it was found that the r.m.s. particle velocity was 1·4 cm. per second, and varied between 0·24 and 2·1 cm. per second for the frequencies $n/3$ to $n/15$. The amplitudes corresponding to these figures ranged between 2×10^{-5} and 7×10^{-4} cm. These amplitudes are small and were presumably smaller still in the narrower tubes, and it seems difficult to associate disturbances of such an order with the effects observed.

---

* 'Proc. Roy. Inst.,' vol. 26, p. 568 (1929–31).

To sum up, the Helmholtz-Kirchhoff equation receives strong support from the present results as regards the variation of tube velocity with radius and frequency. Equation (2) connecting c with the thermal constants of the gas, gives, however, results for c which though correct in order of magnitude are found, when used in the Helmholtz-Kirchhoff equation, to require a modification, taking account of the properties of the tube surface. To a first approximation our observations suggest that this modification can take the simple form of a multiplying factor which amounts to a reduction of about 10 per cent. in the theoretical value of c for smooth-walled tubes, and an increase of about 30 per cent. for a rough surface such as a carbon tube provides. Such a method of expressing tube correction was also used by Dixon, Campbell and Parker,* and by Partington and Shilling.†

Free-Space Velocities of Sound.

We have set out in Table VI the free-space values of the velocity of sound for the temperature range 0° to 100° C. based on the work of a number of recent observers, all of whom used the resonating-tube method, with the exception of Ladenberg and Angerer who worked in the open air. We have considered it justifiable to reduce to 18° C. those values which are reported by the various authors at temperatures from 15° to 20° C. Extrapolation to 0° C. has been made on the assumption that γ, the ratio of specific heats, varies in a linear manner with temperature, the rate of variation being based in each case on the observer's own results. This extrapolation is in great measure independent of the particular equation of state which is assumed to hold for the gas in question, although the actual values of γ themselves vary very considerably according to the correction made for deviation from the laws of an ideal gas. All cases of extrapolation by us which extend over more than 2 or 3 degrees are indicated by enclosing the values in brackets.

The Validity of the Helmholtz-Kirchhoff Formula.

As remarked above, the Helmholtz-Kirchhoff formula has been subjected to a good deal of criticism both by earlier workers and more recently.‡ A review

---

* 'Proc. Roy. Soc.,' A, vol. 100, p. 1 (1921).

† "The Specific Heats of Gases," p. 53 (London).

‡ *Cf.* Partington and Shilling, "The Specific Heats of Gases," p. 53 (1924), and Barton "Text Book on Sound," pp. 542–550 (1908).

Table VI.—Free-Space Velocities in Dry Pure Gases.

Gas.	Velocity (metres/sec) at—			Observer and reference.
	0° C.	18° C	100° C.	
Air (free from CO_2)	331·8	342·6	387·5	Dixon, Campbell and Parker, *loc. cit.*
	331·6	—	—	Gruneisen and Merkel, 'Ann. Physik,' vol. 66, p. 344 (1921).
	330·8	—	—	Ladenberg and Angerer, 'Ann Physik,' vol 66, p 293 (1921).
	(329·8)	342·0	(389·4)	Dixon and Greenwood, 'Proc. Roy. Soc.,' A, vol. 105, p. 199 (1924).
	(331·0)	341·8	387·2	Shilling and Partington, 'Phil. Mag.,' vol 6, p. 920 (1928).
	—	(342·2)	—	Cornish and Eastman, 'J. Amer. Chem. Soc.,' vol 50, p. 627 (1928).
	(331·6)	342·4	387·3	Kaye and Sherratt, this paper (1933).
Hydrogen	1261	—	—	Gruneisen and Merkel, *loc. cit.*
	1261	—	1468	Cornish and Eastman, *loc. cit.*
	(1261)	1301	1462	Kaye and Sherratt (1933).
Carbon dioxide	258·3	—	—	Dixon, Campbell and Parker, *loc. cit.*
	—	266·1	—	Partington and Howe, 'Proc. Roy. Soc.,' A, vol. 100, p 27 (1921).
	258·8	—	—	Tornau, 'Z. Physik,' vol. 12, p 48 (1922)
	(257·7)	266·0	299·9	Shilling, 'Phil Mag.,' vol. 3, p. 273 (1927).
	(258·2)	265·8	297·2	Kaye and Sherratt (1933).
Sulphur dioxide	(209·1)	216·2	244·2	Kaye and Sherratt (1933).
Ammonia	(415·4)	428·8	(485·3)	Dixon and Greenwood, *loc. cit.*
	(415·0)	428·2	481·9	Kaye and Sherratt (1933).
Ethyl chloride	(197·0)	203·8	230·6	Kaye and Sherratt (1933).

of the evidence leads us to believe that much of the criticism has little justification.

The earlier work has been summarized by Auerbach* and by Lübcke in Trendelenburg's Akustik† and need not be dealt with here. We need only refer to the more recent work of Grüneisen and Merkel‡ and Cornish and Eastman,§ who studied air and hydrogen using smooth brass resonators of different lengths and diameters and valve-oscillator methods of frequency production. Grüneisen and Merkel agreed with the form of the Helmholtz-Kirchhoff equation but obtained a value for *c* some 10 per cent. less than the theoretical, which is in accord with the present authors' results for smooth tubes. The results of

---

* Auerbach, "Akustik," Winkelmann's "Handbuch der Physik," vol. 2, 2nd edition (1909).

† Geiger and Scheel's "Handbuch der Physik," vol. 8 (1927).

‡ 'Ann. Physik,' vol. 66, p. 344 (1921).

§ 'J. Amer. Chem. Soc.,' vol. 50, p. 627 (1928).

Cornish and Eastman also support the form of the Helmholtz-Kirchhoff equation, though no values of c are given in the paper.

The evidence against the validity of the Helmholtz-Kirchhoff equation rests in fact mainly on the work of the earlier observers, whose accuracy, especially in the matter of frequency measurement and control, was of necessity much lower than can be obtained nowadays. We may conclude that in general the Helmholtz-Kirchhoff equation is substantially valid for smooth tubes.

Acknowledgments.

We desire to acknowledge our indebtedness to the constructive and helpful suggestions of Mr. J. H. Awbery and to thank Mr. N. Fleming for his co-operation in the amplitude determinations.

Summary.

The velocity of sound at 18° C. and 100° C. has been determined by the resonating-tube method in six different gases contained in six tubes of adjustable length and of different diameters and materials (glass, copper and carbon). A series of frequencies from 500 to 27,000 cycles per second were provided by a valve-oscillator system controlled by a vibrating quartz crystal which also served as a resonance detector.

The results indicate that for all the gases the Helmholtz-Kirchhoff formula is quantitatively correct in its statement of the influence of tube diameter and frequency on the velocity. The formula does not, however, take cognizance of the influence of the wall surface. In the smooth tubes the reduction in velocity below the free-space value was on the average about 10 per cent. less and in the rough tubes about 30 per cent. more than the formula would indicate. An examination of the literature indicates that much of the criticism to which the Helmholtz-Kirchhoff formula has been subjected in the past was unjustified.

The following free-space velocities for the pure dry gases have been deduced from the observations. The probable accuracy is of the order of 1 in 1000 except for hydrogen where it may be a little less.

Free-Space Velocity of Sound (metres per second).

Gas.	18° C.	100° C
Air (CO_2 free)	342·4	387 3
Hydrogen	1301	1463
Carbon dioxide	265 8	297 2
Sulphur dioxide	216 2	244 2
Ammonia	428·2	481 9
Ethyl chloride	203 8	230 6

Conduction of Heat in Powders.

By W. G. KANNULUIK, B.Sc., Fred Knight Scholar, and L. H. MARTIN, Ph.D.,
Melbourne University.

(Communicated by T. H Laby, F.R S —Received January 30, 1933)

Introduction.

While regular use is being made in industry of the thermal insulating properties of powders and other fibrous or cellular materials, the physical principles underlying the processes of heat transfer through such bodies have not received much attention.

The problem, in regard to powders, has been studied by Smoluchowski* in 1911, and later by Aberdeen and Laby.† Smoluchowski's investigation included the effect of such factors as the grain size, the kind, and density of the gas on the conductivity of powders of different materials. The experimental procedure followed in his work is unfortunately open to serious objections, and in consequence the data obtained by Smoluchowski are probably not reliable. Further the method, a cooling one, could not be applied to powders having high conductivities. In this laboratory, Aberdeen and Laby have studied the conductivity of the insulating powder silox filled with several gases, as a function of the gas pressure. The powder was contained between two concentric cylinders and the heat transfer was measured across a central portion where the flow was shown to be strictly radial. Platinum thermometry was used so that reliable data are to be expected. Their experiments, however, were limited to a single powder, and, at the suggestion of Professor Laby, a more extensive investigation using powders of different materials and of various grain sizes has been undertaken.

The essential phenomenon is one of gas conduction in cells, the linear dimensions of which may be (1) large compared to the mean free path of the gas molecules, or (2) with fine powders of the same order of magnitude. In this paper we have attempted to throw some light on the physical processes involved.

---

* 'Bull. Int. Acad. Sci. Cracovie,' Ser. A, p. 129 (1910).

† 'Proc. Roy. Soc.,' A, vol. 113, p. 459 (1926).

In the present work we have used a hot wire method. The conductivity of the powder is measured relative to that of a silver wire, the thermal and electrical constants* of which have already been studied by one of us between $-183°$ C. and 100° C. While on general grounds, absolute methods are preferable the method outlined below is so convenient and rapid in amassing a large amount of data that the above objection was outweighed by the obvious advantages of the method. Satisfactory checks on the reliability of the method can be made by using it to find the conductivity of gases such as air and hydrogen.

Method.

A silver wire is mounted axially in a powder-filled glass tube which can be evacuated. The wire is heated by passing a known current through it, the heat generated being conducted along the wire and through the ends, and laterally through the powder. The heat is then dissipated in an ice and water bath in which the apparatus is totally immersed. Provided the thermal conductivity of the wire is known, the magnitude of the lateral flow of heat through the powder can be calculated from the electrical energy supplied to the wire, and the change in resistance of the wire itself

A thin glass tube, 16 cm. long and of internal diameter 1·00 cm , was closed at each end with thin plane copper caps, the seals being made with picein The silver wire was approximately 0·6 mm. in diameter, and after being stretched along the axis of the tube was soft soldered into each cap. Several different wires were used during the course of the experiments. The tube was filled with, or emptied of, a powder by removing a small screw in one of the caps.

All the electrical quantities were measured with a three dial pattern potentiometer. The requisite measurements are the resistance R_0 of the wire at 0° C., $(\bar{R} - R_0)$ the change in resistance when a heating current of I amp. is passed through the wire. This latter is determined in terms of a standard resistance of 0·01 ohm.

A satisfactory feature of the method is the rapidity with which the data can be obtained and the general reproducibility of the results. The accuracy of the electrical measurements may best be judged from the curve given for hydrogen gas, fig. 1, for which the difference $(\bar{R} - R_0)$ corresponded to a mean rise of temperature of only 1·5° C.

---

* Kannuluik, p. 159.

Theory.

Let λ be the thermal conductivity of the wire, k that of the powder and $\rho_0(1+\alpha\theta)$ the resistance per unit length of the wire at the temperature θ°. The condition holding at the surface of the wire is

$$\pi b^2\lambda\frac{\partial^2\theta}{\partial z^2}+2\pi bk\left.\frac{\partial\theta}{\partial r}\right|_{r=b}+\frac{I^2\rho_0(1+\alpha\theta)}{J}=0, \tag{1}$$

where b is the radius of the wire and J the electrical equivalent of heat.

The first two terms multiplied by dz being the net rate of inflow of heat into the element dz along the wire and over the cylindrical surface, and the last term multiplied by dz being the rate of generation of heat in dz

If the flow of heat from the wire to the walls of the tube is a radial one, (1) simplifies to

$$\pi b^2\lambda\frac{\partial^2\theta}{\partial z^2}+2\pi bh\theta+\frac{I^2\rho_0(1+\alpha\theta)}{J}=0. \tag{2}$$

The quantity h is usually called the ' external conductivity." It is easy to show that it is related to k by the relation

$$h=\frac{k}{b\log_e a/b} \tag{3}$$

where a is the external radius of the glass tube The quantity $b\log_e a/b$ may be called the "space factor" of the tube

The solution of (2) has already been given * It is

$$\left(\frac{1}{\beta l}\right)^2\left(1-\frac{\tanh\beta l}{\beta l}\right)=\frac{2\pi b^2J\lambda(\bar{R}-R_0)}{R_0^2I^2\alpha l}$$

with

$$\beta^2=\frac{2h}{b\lambda}-\frac{I^2R_0\alpha}{2\pi b^2J\lambda l} \tag{4}$$

By solving (4) for β, the value of h can be obtained, and then k deduced from relation (3). For the solving of (4) it is convenient to draw up a table of the function

$$\left(\frac{1}{\beta l}\right)^2\left(1-\frac{\tanh\beta l}{\beta l}\right)$$

and to employ interpolation.

---

* Kannuluik, ' Proc. Roy. Soc.,' A, vol. 131, p. 320 (1931).

As stated above in the foregoing simple solution of the problem a radial flow of heat is assumed, but actually the lines of flow of heat are curved. To obtain a rigorous solution it is necessary to proceed from the general equation (1) and not from (2). We are indebted to Professor T. M. Cherry, of the University of Melbourne, for having carried out this calculation for us. As the solution of (1) we have

$$\theta = \sum_n C_n \left[\frac{I_0\left(\frac{n\pi r}{2l}\right)}{I_0\left(\frac{n\pi a}{2l}\right)} - \frac{K_0\left(\frac{n\pi r}{2l}\right)}{K_0\left(\frac{n\pi a}{2l}\right)} \right] \cos \frac{n\pi z}{2l}, \quad n = 1, 3, 5 .$$

and so

$$\left.\frac{\partial\theta}{\partial z}\right|_{r=b} = \sum_n C_n \frac{n\pi}{2l} \left[\frac{I_1\left(\frac{n\pi b}{2l}\right)}{I_0\left(\frac{n\pi a}{2l}\right)} - \frac{K_1\left(\frac{n\pi b}{2l}\right)}{K_0\left(\frac{n\pi a}{2l}\right)} \right] \cos \frac{n\pi z}{2l},$$

$$\equiv \sum_n C_n \frac{n\pi}{2l} N_{1n} \cos \frac{n\pi z}{2l}, \text{ say,}$$

and

$$\left.\frac{\partial^2\theta}{\partial z^2}\right|_{r=b} = -\sum_n C_n \left(\frac{n\pi}{2l}\right)^2 \left[\frac{I_0\left(\frac{n\pi b}{2l}\right)}{I_0\left(\frac{n\pi a}{2l}\right)} - \frac{K_0\left(\frac{n\pi b}{2l}\right)}{K_0\left(\frac{n\pi a}{2l}\right)} \right] \cos \frac{n\pi z}{2l},$$

$$\equiv -\sum_n C_n \left(\frac{n\pi}{2l}\right)^2 N_{0n} \cos \frac{n\pi z}{2l}, \text{ say.}$$

Also for $-l < z < l$

$$\frac{I^2\rho_0}{J} = \frac{4I^2\rho_0}{J\pi}\left[\cos\frac{\pi z}{2l} - \tfrac{1}{3}\cos\frac{3\pi z}{2l} + \tfrac{1}{5}\cos\frac{5\pi z}{2l} \pm .\right.$$

Hence on substituting in (1) and equating coefficients of $n\pi z/2l$, we get

$$C_n = \pm \frac{\dfrac{4I^2\rho_0}{nJ\pi}}{\pi b^2\lambda\left(\dfrac{n\pi}{2l}\right)^2 N_{0n} - 2\pi bk\dfrac{n\pi}{2l}N_{1n} - \dfrac{I^2\rho_0\alpha}{J}N_{0n}},$$

the sign being positive for $n = 1, 5, 9 \ldots$ and negative for $n = 3, 7, 11 \ldots.$ The resistance of the whole wire is

$$\bar{R} = \int_{-l}^{+l} \rho_0 (1 + \alpha\theta)\, dz \equiv R_0 + \frac{2R_0\alpha}{\pi}\left[C_1N_{01} - \tfrac{1}{3}C_3N_{03} + \tfrac{1}{5}C_5N_{05} - \ldots\right]. \qquad (5)$$

This formula (5) replaces (4) of the simple theory. The conductivity k of the powder is involved in the expression for C_n and is the only unknown. The quantities N_{0n} and N_{1n} depend only on the dimensions of the wire and of the tube and the numerical values can be calculated from the expansions of the functions I_0, I_1, K_0, and K_1.*

A comparison of the rigorous formula (5) with the simple radial flow formula (4) for values of k between 6×10^{-5} and 250×10^{-5} cal. cm.$^{-1}$ sec.$^{-1}$ deg.$^{-1}$ discloses that the simple formula gives results which are approximately 1 per cent. too high. In order to carry out the radial flow correction, it is expedient to get a value of k from the simple formula (4) and then to employ (5) to correct this value by the method of successive approximations. As the correction is a small one, it has not been applied to the powder data. The values of k for air and hydrogen are, however, corrected.

Gas Conduction

Although the simple apparatus used in this work was not designed with the intention of using it with gases, it was thought desirable to get a check on the foregoing theory by measuring k for air and for hydrogen. We wished, also, to obtain the form of the k, p (pressure) curves in order to compare these curves for gases with those obtained for powders.

In fig. 1 the data of hydrogen and air are represented graphically. It will be seen from these curves that convection is almost completely absent over a considerable range of pressures The same result was obtained by Sophus Weber† who used a hot wire method under similar conditions. The value of k for air at 0° C. after applying the radial flow correction is $5 \cdot 85 \times 10^{-5}$ cal. cm,$^{-1}$ sec.$^{-1}$ deg.$^{-1}$ and is in agreement with values found by other experimenters using other variants of the hot wire method The similarly corrected value for hydrogen is $38 \cdot 0 \times 10^{-5}$ cal. cm $^{-1}$ sec. $^{-1}$ deg $^{-1}$ (commercial hydrogen having a purity better than 99 per cent. was used). This value is in good agreement with the mean value given in the International Critical Tables.

The position at present is that the hot wire methods give systematically higher values of the conductivity of air than other methods among which the most fundamental is the "parallel plate" method. Measurements of gas conductivities using an improved hot wire apparatus are now being made, and by varying the experimental conditions it is hoped to discover if any sources of error in the "hot wire" method have been overlooked. For the time being the above values of k for air and hydrogen are regarded as tentative only, but we believe that they do serve to show that the method is a sound one.

---

* Whittaker and Watson, "Modern Analysis," chap. 17.

† 'Ann. Physik,' vol. 54, pp. 325, 437 (1918).

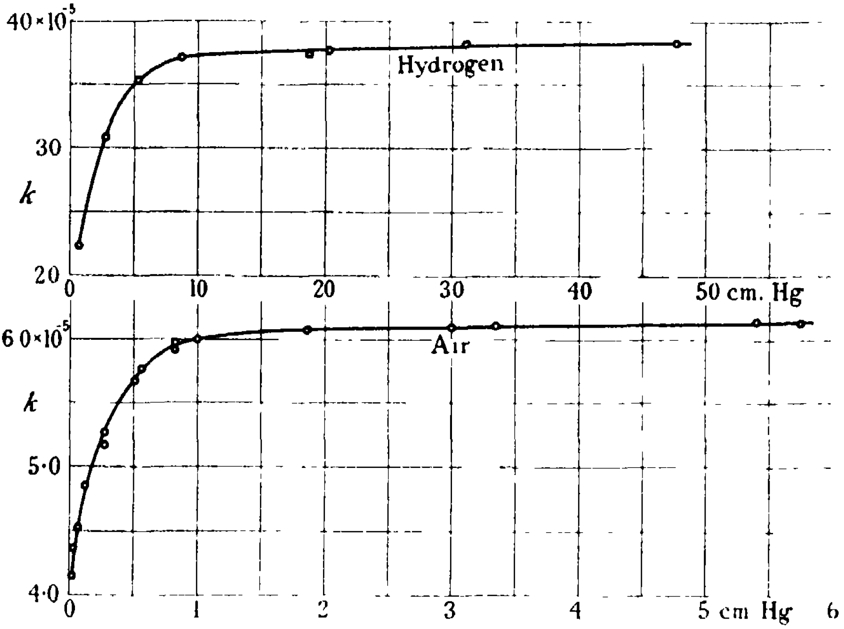


FIG. 1—Conductivity of gases as a function of the pressure.

Experimental Results.

The following is a typical example of the calculations involved in deriving k:—

CO_2 in MgO.

$2a = 1\cdot002$ cm.; $2b = 0\cdot06095$ cm.; $2l = 16\cdot00$ cm;
$\lambda = 0\cdot999$ cal. cm.$^{-1}$ sec.$^{-1}$ deg.$^{-1}$, $\alpha = 0\cdot00400$; $J = 4\cdot18$ watt-sec. cal^{-1}.
$R_0 = 0\cdot0080123_8$ ohm; $\bar{R} - R_0 = 0\cdot00011925$ ohm,
$I = 4\cdot0016_8$ amp.

Substituting in (4) we get

$$\left(\frac{1}{\beta l}\right)^2\left(1-\frac{\tanh \beta l}{\beta l}\right)=0\cdot084781.$$

This is satisfied by $\beta l = 2\cdot7455$, hence $\beta^2 = 0\cdot11778$. From the second relation of (4) we get

$$0\cdot11778=\frac{2h}{b\lambda}-\frac{I^2R_0\alpha}{2\pi b^2J\lambda l}=\frac{2h}{b\lambda}-0\cdot00263,$$

Table I.—Conductivity of Powders.

p, cm. Hg. k, 10^{-5} cal sec.$^{-1}$ deg.$^{-1}$ cm.$^{-1}$.

	Magnesium oxide.			Carborundum.														Glass.		Diphenylamine.	
				No 40.			No. 90.				No 280.			No. 600.						
	H_2.	Air	CO_2	H_2.	Air	CO_2	H_2	Air	CO_2	He	H_2	Air	CO_2.	H_2	Air	CO_2	He	Air	H_2.	Air.
k	46·6	17·5	14·9	253	55·5	43·4	307	62·2	44·5	236	171	62·9	43·5	128	43·2	32·7	86·2	38·3	45·8	19·1
p	76·2	75·7	72·8	76·1	76·0	75·0	71·2	76·4	73·3	63·3	75·7	74·8	70·4	76·0	75·4	76·9	57·4	76·3	76·0	76·5
k	44·8	17·0	14·7	247	54·8	43·5	277	61·4	44·1	214	157	59·2	41·6	115	40·5	32·0	85·4	38·6	44·6	19·0
p	62·5	57·5	62·0	57·0	59·0	60·6	40·8	53·4	56·6	41·7	57·6	55·4	52·3	57·3	58·0	59·7	39·1	64·2	37·4	57·5
k	42·7	16·2	14·4	236	53·8	41·5	251	58·8	43·3	202	142	56·1	40·5	95·5	37·5	29·4	83·4	38·1	44·1	18·7
p	50·3	40·4	48·2	34·5	44·2	41·1	27·1	33·1	41·2	34·8	40·6	38·3	37·0	36·1	43·4	40·1	27·9	52·4	26·5	46·6
k	40·3	15·3	13·6	224	52·5	40·9	212	51·9	42·2	169	119	49·6	36·8	72·3	32·3	24·2	80·0	37·7	43·1	18·5
p	39·3	26·8	29·8	21·2	26·5	25·9	14·5	13·3	28·6	19·2	25·2	22·7	22·4	19·7	25·4	20·5	18·6	40·3	16·7	32·4
k	36·8	14·2	13·0	197	49·8	38·5	174	43·9	40·4	131	86·0	39·8	32·8	43·5	21·8	19·5	74·7	37·2	39·8	18·0
p	27·3	17·7	21·3	13·1	13·7	12·9	9·74	5·23	16·6	10·3	12·4	9·65	11·6	8·17	9·6	10·1	11·3	30·3	7·46	15·8
k	31·0	13·7	12·1	186	45·4	36·5	143	39·0	35·8	96·6	60·7	31·7	25·0	23·9	11·8	14·8	69·5	36·6	31·2	17·7
p	15·1	14·1	13·9	8·26	6·29	7·39	6·04	3·45	7·62	5·45	6·46	4·35	4·29	3·01	2·40	5·06	7·14	19·8	1·78	7·22
k	27·6	12·6	11·6	168	42·2	31·2	99·1	20·5	32·1	63·3	32·1	16·8	19·8	12·9	7·34	11·6	58·1	34·1	20·3	16·3
p	10·4	10·0	9·73	5·32	4·08	2·66	2·54	0·443	3·84	2·52	2·19	0·98	2·18	0·931	0·718	2·78	3·52	9·97	0·49	2·55
k	22·6	11·9	9·98	146	37·1	26·7	35·1	14·0	24·6	30·6	11·0	13·9	13·9	8·15	5·03	7·93	36·1	30·2	—	14·5
p	5·72	8·20	5·15	3·39	2·22	1·25	0·62	0·102	1·34	0·74	0·455	0·58	0·76	0·244	0·146	1·11	1·00	3·24	—	1·04
k	14·9	11·2	—	116	32·1	—	—	—	—	—	—	11·8	—	—	—	5·39	—	—	—	—
p	1·88	6·16	—	1·96	1·14	—	—	—	—	—	—	0·356	—	—	—	0·32	—	—	—	—

finally

$$k = 2{\cdot}303b \log_{10} a/b \times h$$
$$= 15{\cdot}7 \times 10^{-5} \text{ cal. cm.}^{-1} \text{ sec.}^{-1} \text{ deg.}^{-1}.$$

Table I gives the collected values of k for the different powders investigated. The relevant particle dimensions and densities are shown in Table II.

Table II.—Densities of powders (ρ), materials (ρ_0), and linear dimensions of particles (l)

	Carborundum.				Magnesium oxide	Glass	Diphenyl-amine.
	No. 40.	No 90	No. 280	No. 600			
ρ, g cm^{-3}	1·83	1·89	1·84	1·54	0·20	1·94	0·565
ρ_0, g cm^{-3}	3·20	3·20	3·20	3·20	3·08	2·98	1·16
l, cm	0·055	0·0194	0·0061	0·0027	0·001	0·032	0·0027

General Considerations of the Data.

The experimental results can be represented graphically by plotting k either directly against the pressure p or against $\log p$. The latter alternative is adopted by both Smoluchowski and by Aberdeen and Laby, and has the advantage of giving a better spacing of the results at low pressures. For their results Aberdeen and Laby obtained a linear relationship between k and $\log p$. The dotted curve in fig. 2 which gives their results for hydrogen in silox shows that if full weight be given to each of their points, a S-shaped curve is obtained. This, we have shown to be the typical form of k, $\log p$ curves for gas-filled powders. For example, fig. 2 contains representative k, $\log p$ curves for air-filled carborundum powders of different grain sizes obtained by us. The general form of these curves is an S-shaped one, but within the range of pressures used by us the complete S is not normally obtained. Powder No. 90, however, shows a typical complete S.

The general features of heat conduction in powders are brought out more clearly by the k–p curves than by the k, $\log p$ curves. The k, p curves are shown in fig. 3 for the four grades of carborundum powder. The group for carbon dioxide shows most completely the variation of k with p, at high pressures the conductivities tend to limiting values, while at low pressures the curves bend sharply downwards so that k falls rapidly with decreasing pressure. The air and hydrogen curves show in a partial manner the same features as the carbon dioxide curves. It is as though the scale of the pressure were successively magnified in passing from carbon dioxide, to air, to hydrogen. In general for a given pressure the finer the powder the further k is from the limiting value. This feature is of considerable significance and should be a guide in choosing suitable powders for thermal insulators.

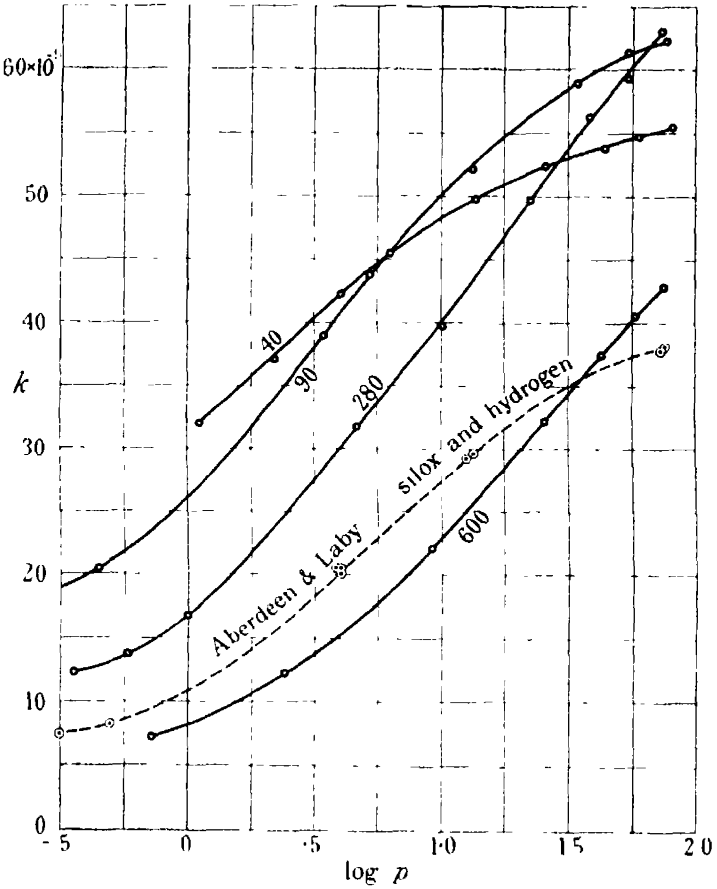


Fig. 2.—k, $\log p$ curves for air-filled carborundum powders and silox in hydrogen.

The curves, particularly those of the carbon dioxide group, resemble the adsorption isotherms for gases on solids, and, at first sight, it might be thought that adsorbed gas was forming a conducting layer on the surface of the grains, and so causing a breakdown of the thermal resistance between the proximate surfaces of the grains. Experiments on glass powder indicate that this explanation is improbable. On evacuating and then filling this powder with air at atmospheric pressure equilibrium is reached only after a lapse of some days, and yet the conductivity is of the same order of magnitude as was observed for a carborundum powder of similar grain size in which no adsorption could be observed. Some of the data for the glass powder were obtained before the air had reached the equilibrium pressure and the pressure was still

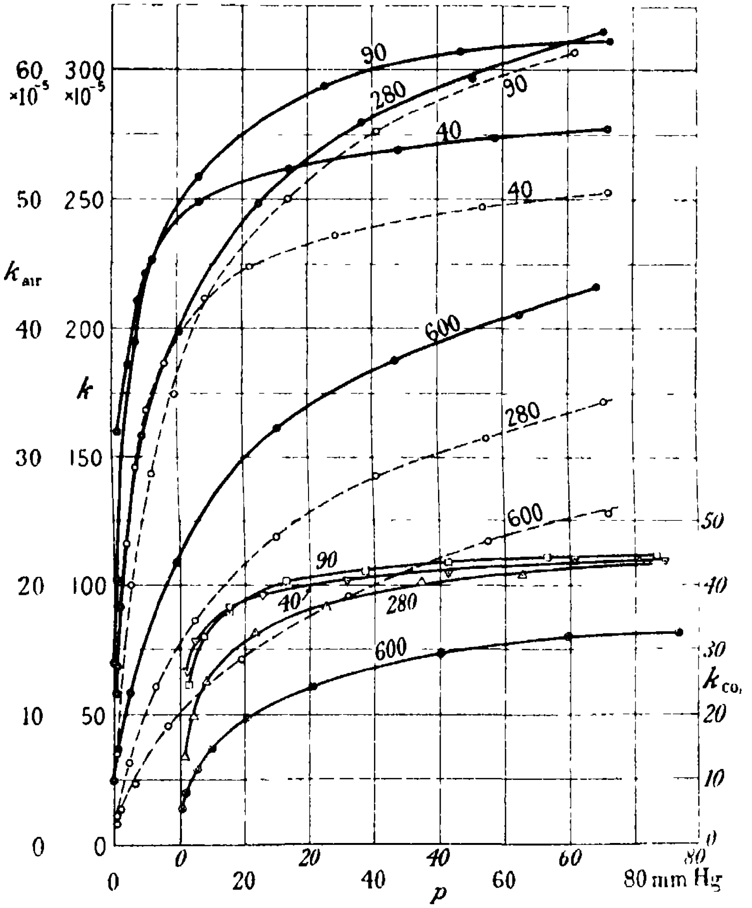


FIG. 3.—k, p curves for carborundum with H_2, shown in dotted line, and filled with air and carbon dioxide, solid line.

slowly changing, yet the conductivities fitted satisfactorily on the typical curve. Finally, an experiment was made in which this powder was filled with helium which showed no adsorption. Table I shows the high conductivity which was obtained, indicating that adsorption as a possible explanation of the phenomena is unlikely.

That the conductivities of the powder and the gas are not simply additive is evident from the value of the conductivities found for the evacuated powders which are of the order $0 \cdot 1 \times 10^{-5}$. As a typical example the conductivity of a hydrogen-filled carborundum powder was as high as 300×10^{-5}, while that of hydrogen is $38 \cdot 0 \times 10^{-5}$.

Theoretical Considerations.

Before considering the theoretical implications of the data, it is necessary to draw attention to the fact that owing to the finite size of the powder particles and of the gas cells, the lines of heat flow regarded in detail must depart from the ideal mathematical distribution presupposed in calculating a space factor. Also it may be questioned whether there is adequate contact between the wire and the powder and whether this contact varies with the grain size.

With regard to the first it must be admitted that this is an inherent objection to employing the conception of thermal conductivity at all in reference to a powder, and raises the question whether it would not be preferable to treat the problem as one of heat transfer. On the other hand, the traditional treatment has obvious advantages. The second objection, if it were a valid one, would lead to incorrect absolute values of the conductivity, but would not affect the variation of the conductivity with the pressure. A consideration of the data of all our powders does not disclose any inconsistencies in the absolute values of the conductivities obtained. Further, the conductivities of the evacuated powders are all of the same order of magnitude (1×10^{-5}) so that the thermal resistance between the wire and the powder is sensibly the same for both fine and coarse powders.

A theory of conduction in powders has been attempted by Smoluchowski (*loc. cit.*) in which he treated the ideal case of spherical grains with cubical packing. He calculated the heat transfer between adjacent hemispheres taking into account the temperature discontinuity at the solid-gas interface. As a result of this discontinuity the normal gradient $\Delta T/d$ must be replaced by $\Delta T/(d + 2\delta)$ where δ is a length varying inversely with the gas density and is usually designated by the German word "Sprung" He obtained for the powder conductivity the expression

$$k = \frac{\pi k_0}{2} \log\left(1 + \frac{a}{\delta}\right)$$

where k_0 is the conductivity of the gas and a is the radius of the grain spheres. This formula is inadequate to explain our experimental results as it fails to give a limiting value of k. This defect is due essentially to the fact that he considered grains of spherical shape which in the analysis led to the introduction of the log function. Further with this model it was impossible to introduce into the theory the conductivity of the solid phase.

In an attempt to overcome these objections we have assumed an idealized picture of a powder, namely, one which consists of a large number of laminæ of solid separated by layers of gas. Let k, k_γ, and k_s be the conductivities of powder, gas, and solid respectively, then

$$\frac{1}{k} = \frac{\sum_n l}{k_s} + \frac{\sum_n (d + 2\delta),}{k_g} \tag{6}$$

n being the number of layers of solid and gas per centimetre; l the thickness of a solid layer; and d that of a gas layer. Putting

$$\sum_n l = \rho/\rho_0 \equiv \omega \qquad \text{and} \qquad \sum_n d = 1 - \omega,$$

where ρ is the density of the powder ρ_0 that of the solid phase of the powder and $2\delta = c/p$, c being a constant, we get

$$k = \frac{k_g}{1 - \omega\,(1 - k_g/k_s) + \dfrac{c\,(1 - \omega)}{d \,.\, p}} \tag{7}$$

when p becomes large k approaches a limiting value given by

$$k = \frac{k_g}{1 - \omega\,(1 - k_g/k_s)}\,. \tag{8}$$

Our k, p curves show or at least indicate the reality of the limiting value of k. A numerical test of (8), however, does not lead to wholly satisfactory results. For example, for glass powder filled with helium it gives at atmospheric pressure a limiting value of $k = 73 \times 10^{-5}$, while filled with air the limiting value at this pressure is 16×10^{-5}. The experimental values are 88×10^{-5} and 38×10^{-5} respectively. An attempt is made to explain these differences later.

Equation (8) shows that the limiting value of k is determined by the packing factor ω and the ratio k_g/k_s. As ω may be comparable with unity the value of k can be large compared with conductivity k_g of the gas. For the limiting value of k to be less than that of the gas, k_g must be greater than k_s. For diphenylamine,* which has a conductivity (5×10^{-4}) of the same order of magnitude as that of hydrogen, a limiting value approaching the conductivity of hydrogen should be obtained by using a loosely packed powder. We verified this conclusion experimentally. (See data in Table I for diphenylamine.)

---

* Lees, 'Phil. Trans.,' A, vol. 204, p 433 (1905).

With coarse powders it was found that (7) represents adequately the variation of k with p, as will be seen by rewriting it in the form

$$p/k = p\left|\frac{\omega}{k_s} + \frac{1-\omega}{k_g}\right] + \frac{c(1-\omega)}{d\,.\,k_g}. \tag{9}$$

When p/k was plotted against p for the coarse powders, diphenylamine with air and hydrogen, and glass with air and helium it was found to be a linear function of p. But with the hydrogen filled carborundum powders the linear function breaks down for the finer grades Nos. 280 and 600.

Apart from the fact that the powder model used above falls far short of representing the actual conditions for a powder, there is a process of heat transfer the existence of which has been ignored in the simple theory. It has already been mentioned that the conductivities of the evacuated powders were extremely small, being of the order 1×10^{-5}. This is due, of course, to the small areas actually in contact between contiguous grains. There must, however, be large surfaces whose distances apart are less than the mean free path of the gas molecules, and in these gas layers heat will be transferred by molecular conduction. The finer the powder and the closer the packing the greater will be the contribution to the heat flow by this process. The failure of the simple theory outlined above should be in greater evidence in the hydrogen-filled powders than in those filled with air and carbon dioxide since the mean free path of the hydrogen molecules is nearly double that of these latter gases at a given pressure. This was found to be so by experiment.

Our previous equations must now be altered to include the effect of layers of gas for which there is molecular conduction. If we assume $n - m$ gas layers per centimetre each of thickness s for which there is molecular conduction (6) takes the form

$$\frac{1}{k} = \frac{\sum_n l}{k_s} + \frac{\sum_m (d + 2\delta)}{k_g} + \frac{1}{p\varepsilon}, \tag{10}$$

where ε is a constant given according to Knudsen* by

$$Q = \varepsilon\,.\,p\,.\,\Delta\theta,$$

---

* Lorentz, "Theoretical Physics" (Macmillan), vol. 1, p. 136.

where Q is the heat transfer per second between two plates with a temperature difference $\Delta\theta$ the distance apart of which is less than the mean free path. It has been implicitly assumed that $(n - m)\,s$ is independent of the pressure and the linearity of p/k, p curves for glass and diphenylamine justifies this assumption for coarse powders, at least for the higher pressures. Since the number of gas layers of thickness d has been reduced from n to m, it follows from (10) that the values of k calculated from this must be greater than those calculated previously from (8) so that (10) represents an improvement as regards absolute values of k. Equation (10) can be written as follows

$$p/k = p\left\{\frac{\omega}{k_s} + \frac{1 - \omega}{k_g} - \frac{(n - m)\,s}{k_g}\right\} + \frac{1}{d\,.\,k_g}\{c\,(1 - \omega) - (n - m)\,s\} + \frac{1}{\varepsilon} \quad (11)$$

This is of the same form as (9). A quantitative comparison of equation (11) with experiment presents obvious difficulties. However, it is possible to evaluate $(n - m)\,s$ from the slopes of the p/k, p curves and then to substitute the values in the residuals and so obtain numerical values of d. For glass powder we obtained 0·023 mm. and 0·059 mm for air and hydrogen respectively, while for diphenylamine the values were 0·028 mm. and 0·034 mm. for air and helium. These values are of the right order of magnitude and as would be expected the value of d increases for a given powder as the molecular weights of the gas decreases.

For the lowest pressures in these two powders the experimental points fall below the straight line $p/k = ap + b$, while for the two finest carborundum powders the departure sets in at quite high pressures. This results from high values of k and presumably these are due to the fact that as the pressure decreases an increasing proportion of the gas layers conduct heat according to the molecular process. To account for this theoretically it would be necessary to modify (11) by making $(n - m)$ a function of the pressure. In view of the multiplicity of adjustable constants, there is little to be gained by developing the theory any further along these lines. We do, however, believe that it is necessary to include both processes of conduction in order to account for the experimental results.

In conclusion one of us (W.G.K.) is indebted to the University of Melbourne for the award of a Fred Knight Research Scholarship. It is a pleasure to express our thanks to Professor T. H. Laby, F.R.S., for the interest he has taken in this work, and for helpful criticisms in writing up this paper.

Summary.

A description is given of a hot wire method of determining the thermal conductivity of powders in gases As powders the following materials were used · glass spheres, diphenylamine, magnesium oxide, and a series of graded carborundum powders. These materials were immersed in turn in air, hydrogen, carbon dioxide and in some cases helium

The conductivities were measured over a range of pressures 0·5 cm. to 76 cm

It is found that the conductivity k, can be expressed as a function of the pressure p by means of the relation

$$p/k = ap + b$$

A simple derivation of this relation is given making use of the Fourier equation, and the temperature discontinuity at the gas-solid interfaces (temperature "Sprung")

Departures from this relation observed for some of the finer powders, and in general for all powders at low pressures, are explained by introducing the conception of molecular conduction

The Thermal and Electrical Conductivities of Several Metals between — 183° C. and 100° C.

By W. G. KANNULUIK, B.Sc., Fred Knight Scholar, University of Melbourne.

(Communicated by T. H. Laby, F R S.—Received January 30, 1933)

Introduction.

In a previous paper* an electrical method of determining the thermal conductivity of metals in wire form was described. The apparatus was of a simple kind, the wire being soldered axially in a glass tube with metal end-pieces. By maintaining a vacuum of the order of 10^{-4} mm. Hg in the tube, the heat developed in the wire on passing a steady current through it is conducted along it and through its ends except for a small amount lost laterally to the walls of the tube by molecular conduction and by radiation The thermal conductivity of the wire is calculated from its known dimensions, the rate at which electrical energy is supplied to it and the change in its electrical resistance. A small correction is applied to allow for the lateral loss of heat.

The method has a number of advantages —

(1) The apparatus used meets the requirements of theory almost exactly and it permits the use of wires of any diameter

(2) The direct determination of temperatures is not required, as they are in effect obtained by measuring changes in the resistance of the wire.

(3) The method lends itself particularly to low temperature work

(4) The electrical conductivity is obtained from observations which are also used in determining the thermal conductivity, so that both are obtained under the same conditions.

The present paper is an account of the application of the method to obtain the thermal and electrical conductivities of pure specimens of tungsten, molybdenum, silver, and iron between —183° C. and 100° C.

Description of the Apparatus.

The two tungsten wires W1 and W2 used in this work are monocrystals with natural crystal faces, the section of W1 being rectangular, that of W2 hexagonal. (The axis of a wire of rectangular section lies in the (100) plane, that of a wire of hexagonal section in the (111) plane.) As these wires are somewhat fragile below 0° C., it is necessary to protect them from breaking at low temperatures, the tube in which they are mounted having the greater thermal expansion. This is done by fixing one end of the wire in a diaphragm of thin sheet copper enclosed in the tube. An internal diaphragm is desirable in order to keep the pressure on both sides of it the same when the tube is evacuated. The use of a tube with an internal diaphragm is practically convenient, and was adopted for the silver and molybdenum wires also. A longitudinal section of the tube is shown in fig. 1. The central portion JJ′ is of lead-glass tube of 1 cm. internal diameter, the rest of the tube being of copper. The copper-glass joins J and J′ are vacuum-tight over the range −183° C. to 100° C. The diaphragm D (2 cm. in diameter) is of sheet copper 1/10 mm. thick and contains the concentric copper stud S (6 mm. in diameter) riveted and gold-soldered to it. The stud provides a base into which to solder one end of the wire and helps in dissipating the heat from this end through the diaphragm to the end of the tube. In considering the effect of the diaphragm, S can be ignored and the diaphragm treated as a simple annulus. In order to evacuate the tube the diaphragm is pierced near its periphery by three spaced holes about 1 mm. in diameter. I, I′ and P, P′ are current and potential leads respectively.

---

* 'Proc. Roy. Soc.,' A, vol. 131, p. 320 (1931).

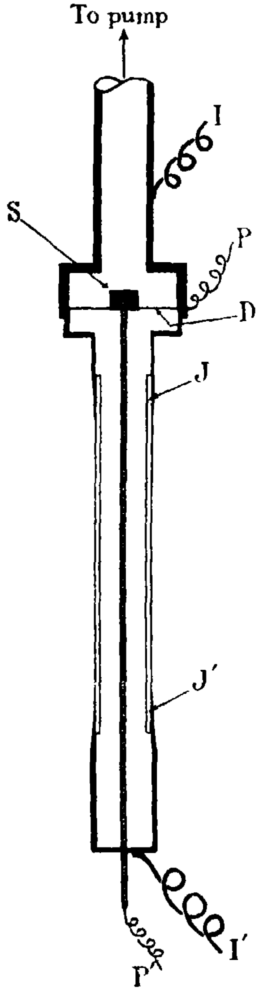


FIG. 1.

When both ends of the wire are strictly at the temperature of the bath in which the tube is immersed, it has been shown (*loc. cit.*) that

$$\lambda = \tfrac{1}{6}\,\frac{RR_0I^2\alpha l}{Jq(R-R_0)}\left(1+\frac{R_0I^2\alpha l}{30J\lambda q}\right)\left(1-\frac{ph}{q\lambda}\,l^2\right), \quad (1)$$

where λ is the thermal conductivity of the wire; R ohm its resistance when it carries a current of I amp.; R_0 its resistance at the temperature of the bath; α its temperature coefficient of resistance; *h* cal. cm.$^{-2}$ sec.$^{-1}$ deg.$^{-1}$ the (Fourier) external conductivity, *i.e.*, the lateral loss of heat from the wire to its surroundings; 2*l* cm., the length of the wire,

q cm.^{2} its cross-section, p cm. its circumference, and J watt cal.$^{-1}$ the electrical equivalent of heat.

Diaphragm Correction.

Owing to the use of the diaphragm in this work, values of λ calculated from (1) have to be multiplied by a small correcting factor. Allowance must be made for the circumstances that the observed resistances of the wire include the resistance ΔR of the diaphragm and that the end of the wire in the diaphragm is at a slightly higher temperature than that of the bath. A simple calculation leads to the following correcting factor:

$$\left(1 - \frac{2\Delta R}{R_0} + \frac{3CJ\lambda q}{l}\right),$$

where $\Delta R = \frac{\sigma}{2\pi t} \log_e a/b$, and $C = \log_e a/b/4\pi J\Lambda t$, a and b being the outer and inner radii of the diaphragm respectively, t its thickness, σ its specific resistance, and Λ its thermal conductivity. The diaphragm correction involves an increase in the values of λ calculated from (1) of the order of ½ per cent. and is approximately the same for the different baths used. It is sufficient to calculate the correction from the dimensions of the diaphragms and the assumed values of the thermal and the electrical conductivity of copper. As a typical example, the details for a silver wire 9·77 cm. long and 0·06096 cm. in diameter are as follows. The resistance of the wire at 0° C. is $R_0 = 0{\cdot}00507563$ ohm, the diaphragm 0·0137 cm. thick has radii of 1·09 and 0·31 cm.; its calculated resistance is $\Delta R = 22{\cdot}5 \times 10^{-6}$ ohm at 0° C. and the value of C is 1·835, the thermal conductivity of copper being assumed to be 0·92 cal. cm.$^{-1}$ sec.$^{-1}$ deg.$^{-1}$. The approximate value of λ for the silver wire being 1·00 cal. cm.$^{-1}$ sec.$^{-1}$ deg.$^{-1}$, we find that

$$1 - \frac{2\Delta R}{R_0} + \frac{3CJ\lambda q}{l} = 1 - 0{\cdot}009 + 0{\cdot}014 = 1{\cdot}005.$$

At the temperatures $-183°$ C., $-78{\cdot}50°$ C. and 100° C., the diaphragm factor has approximately the same value.

The Correction for the Lateral Loss of Heat from the Wire.

The method of allowing for the small lateral loss of heat is the same as that previously outlined (*loc. cit.*, p. 328); the residual gas pressure in the tube is, as in the previous work, less than 10^{-4} mm. Hg. In this work the lateral losses have been reduced by using relatively shorter and thicker wires. Lateral loss of heat takes place by molecular conduction and by radiation. In comparison with the radiative loss that by molecular conduction is small enough to be neglected; the radiative loss at —183° C. can also be ignored. There remains a small correction for radiation at 100° C., 0° C. and —78·5° C. The magnitude of the correction for the different wires is given in Table I, in which the numbers are $\Delta\lambda/\lambda = 2/5\,phl^2/q\lambda$. After applying the diaphragm correction to them the values of λ are to be multiplied by $(1 - \Delta\lambda/\lambda)$.

Table I.—Correction for the Lateral Loss of Heat ($\Delta\lambda/\lambda$).

Metal.	100° C.	0° C	—78·5° C.
Iron	—	0·0090	0·0022
Tungsten—			
W1	0·008	0·0030	0·0011
W2	0·014	0·0053	0·0028
Molybdenum—			
Mo1	0·024	0·0088	0·0032
Mo2	0·014	0·0053	0·0019
Silver	0·003	0·0011	0·0004

Experimental Details.

An account of the experimental procedure and a diagram of the electrical circuit have been included in the previous paper (*loc. cit.*). All the electrical quantities in (1) are obtained by means of a three-dial pattern potentiometer. The constant temperature baths used are the following: Liquid oxygen (B.P. —182·97° C.); solid CO_2 in methylated spirits (S.P. —78·50° C.); ice and water (0° C.); steam (100·000° C.); and naphthalene (B.P. 217·96° C.).

The electrical quantities R_0, R, I, and α are all directly determined except α, the value of which is deduced for any temperature t° C. from the observed variation of R with t over the range —182·97° C. to 217·96° C. To represent this variation the following three and four constant formulæ were employed:

$$\begin{cases} r_t = 1 + at + bt^2 + ct^3 \\ r_t = 1 + at + bt^2 + ct^3 + dt^4, \end{cases}$$

where r_t is the ratio of the observed resistance R_t to the resistance R_0 at 0° C. and a, b, c, and d are constants. Except in one or two cases the three constant formula was used. Using the latter formulæ one gets:

$$\alpha_t = (a + 2bt + 3ct^2 + 4dt^3) \div R_t/R_0.$$

The only other quantity requiring notice is the cross-section q of the wire. As relatively thick wires are used, q can be obtained with considerable precision. It was obtained from measurements of the diameter excepting for the tungsten wires. A Prestwich fluid gauge was used to compare the diameters of the wires with Johansson gauges. This correction was a very small one.

The mean cross-section q of the tungsten monocrystal wires was deduced from the length of the wires and their loss of weight in water. At most these wires are only approximately uniform in cross-section, and so the electrical and thermal conductivities obtained for them cannot be regarded as being as reliable as those obtained for the other wires.

The quantity which limits the reproducibility of λ for the different values of I is the resistance change $(R - R_0)$. Working with a minimum mean rise of temperature of 2½° and with the tube at 0° C., this difference can be obtained to within 1 in 250, provided that R and R_0 are both obtained to within 1 in 50,000. The actual reproducibility of λ for different wires at 0° C. using a range of values of I did not fall below 1 in 250, while a similar accuracy was obtained at the other bath temperatures. It should be noted that the use of the internal diaphragm does not in the least affect the accuracy with which it is possible to get $(R - R_0)$.

The Purity of the Wires.

An important consideration in thermal and in electrical conductivity work (and one which has not always received adequate attention) is the purity of the metal specimens used. As a general rule the electrical conductivity is more sensitive to the effect of impurities than the thermal. The resultant diminution of both conductivities depends not only on the total amount of the impurities present but also on the nature of the impurities themselves.

The writer is indebted to Dr. C. E. Eddy and Mr. T. H. Oddie of this laboratory for making a (quantitative) optical spectroscopic analysis of specimens of three of the wires used in this work, viz., tungsten, molybdenum, and silver. The results of the analyses made by them are contained in the appendix to this paper.

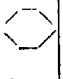
While the writer has used the purest wires that were available to him, the iron wire (by Armco) is hardly as pure as is desirable for this investigation. A chemical analysis kindly supplied by the Armco Corporation gives the following impurities: carbon, 0·011 per cent.; manganese, 0·017 per cent.; phosphorus, 0·006 per cent.; sulphur, 0·026 per cent.; copper, 0·056 per cent.; and silicon, 0·002 per cent. Presumably the iron content is 99·88 per cent.

Experimental Data.

The data obtained for the various wires are given in Table II.

Table II.

R_0 = resistance in ohms at 0° C.; $r_t = R_t/R_0$ = ratio of resistance at t° to that at 0° C.
σ = specific resistance in ohm cm; λ is in cal. cm.$^{-1}$ sec.$^{-1}$ deg.$^{-1}$.
L = $\lambda\sigma/T$ watt ohm deg.$^{-2}$ = Lorenz coefficient.
q cm.^{2} = mean cross-section of the tungsten crystals.

Wire	Length.	Diameter.	R_0	Temperature	r_t.	$\sigma \times 10^6$.	λ.	$\frac{\lambda\sigma}{T} \times 10^8$.
	cm.			°C				
Fe	14·53	0·3924 cm.	0·001150	100	1·6183	15·49	—	—
				0	1	9·57	0·1688	2·47
				−78·50	0·5995	5·74	0·1710	2·11
				−183·00	0·1600	1·531	0·224	1·60
W1	7·846	(q = 0·01053 cm.^{2})	0·003694	100	1·4755	7·35	0·389	3·20
				0	1	4·98	0·399	3·04
				−78·50	0·6464	3·22	0·405	2·80
				−183·00	0·1790	0·892	0·461	1·91
W2	7·940	(q = 0·01022 cm.^{2})	0·003836	100	1·4764	7·29	0·390	3·19
				0	1	4·94	0·405	3·06
				−78·50	0·6424	3·17	0·425	2·90
				−183·00	0·1706	0·843	0·511	2·00
Mo1	12·627	0·09979 cm.	0·008472	100	1·4610	7·67	0·328	2·84
				0	1	5·25	0·327	2·63
				−78·50	0·6465	3·39	0·326	2·38
				−183·00	0·1814	0·952	0·438	1·94
Mo2	9·859	0·09980 cm.	0·006518	217·96	2·0367	10·05	—	—
				100	1·4622	7·56	0·329	2·79
				0	1	5·17	0·329	2·61
				−78·50	0·6434	3·33	0·329	2·35
				−183·00	0·1705	0·882	0·430	1·76
Ag	9·770	0·06095 cm.	0·005053	100	1·4055	2·121	0·996	2·37
				0	1	1·509	0·999	2·31
				−78·50	0·6862	1·036	1·004	2·24
				−183·00	0·2501	0·377	1·015	1·80
Ag*	9·770	0·06095 cm.	0·005056	217·96	1·8959	2·863	—	—
				100	1·4056	2·123	0·998	2·37
				0	1	1·510	0·999	2·31
				−78·50	0·6853	1·035	1·005	2·24
				−183·00	0·2261	0·341	1·018	1·62

* The measurements were repeated after the silver wire had received a prolonged annealing at 500° C.

Discussion of the Results.

Thermal Conductivity.

Tungsten.—As already stated the tungsten data are probably not so reliable as regards their absolute values as those for the other wires, as the tungsten wires were of not quite uniform cross-section. The only other low temperature determinations for tungsten crystals are those of Gruneisen and Goens.* These authors obtained values of λ of 8·20 and 0·555 at −252° C. and −190° C. respectively. The writer's value of 0·511 at −183° C. for W2 accords with the above values of Grüneisen and Goens, when allowance is made for the rapid increase in λ below −183° C.

Molybdenum—The values of λ at 0° C. are slightly smaller than those previously obtained for another wire. But this wire was probably somewhat purer than the wires used in the present work, as it possessed appreciably higher temperature coefficients of resistance. The thermal conductivity at low temperatures does not appear to have been investigated hitherto.

Silver.—The 0° C. value of λ = 0·999 is in close agreement with the value of 1·006 at 18° C. obtained by Jaeger and Diesselhorst† for a rod containing 99·98 per cent. silver. The variation of λ with temperature has been studied by Lees‡ between 0° C. and −170° C., but he used a somewhat impure specimen containing 99·90 per cent. silver. At −170°C. Lees obtained an increase of only 2 per cent. on the value of λ at 0° C., which is the same as that obtained here with a much purer silver. In this respect the behaviour of silver is anomalous; most other metals showing 15 to 20 per cent. increase in λ over this range of temperature.

Iron.—Values of λ at 0° C. by different investigators range from 0·15 to 0·23, principally on account of the great diversity in the degree of purity of the specimens used. Another factor which cannot be ignored is the macroscopic crystal structure of the iron, for Eucken and Dittrich,§ working with an electrolytic iron, found a 10 per cent. greater conductivity at 0° C. for their coarsest as compared with their finest grained specimen. Though the Armco iron studied by the writer is comparatively impure, it is of interest that the variation of λ with temperature between 0°C. and −183° C. is typical of that of a pure metal.

The Lorenz Coefficient.—The high values of the Lorenz coefficient at 0° C.

---

* 'Z. Physik,' vol 44, p. 615 (1927).

† 'Wiss. Abh. Phys.-tech. Rchs-Anst. Berl.,' vol. 3, p. 269 (1900).

‡ 'Phil. Trans.,' A, vol. 208, p. 381 (1908).

§ 'Z. phys. Chem.,' vol. 125, p. 211 (1927).

obtained for the tungsten monocrystals are noteworthy. These values of $3 \cdot 1 \times 10^{-8}$ are much in excess of $2 \cdot 43 \times 10^{-8}$ predicted by the Sommerfeld-Fermi theory for the Lorenz coefficient, and are actually the greatest values which have been observed in a pure metal. Some support is given by these figures for tungsten and also to a lesser extent by those of molybdenum to the contention of Eucken * and his co-workers that the part played by the lattice vibrations in the heat conduction process is real, and must be taken into account before a comparison of the theoretical value with the experimental values of the Lorenz coefficient can be made. There is, however, a difference of opinion as to how this is to be done, but the simple assumption made by Eucken that the thermal conductivity of a metallic conductor is the sum of an "electron" and a "lattice" conductivity has led to useful results in explaining the abnormally high values of the Lorenz coefficient of some of the poorer conductors such as antimony, bismuth, and certain alloys. This treatment of the problem cannot further the cause of theory in the region of low temperatures, where the Lorenz coefficient falls away at −252° C. by as much as two-thirds of its value at 0° C. in a number of pure metals. Recent theoretical work by Peierls† on the thermal conductivity of metals indicates that the Lorenz coefficient of sufficiently pure metals varies as the temperature at low temperatures, but it discounts the part played by the lattice vibrations, more especially in the region of low temperatures. This theory is, moreover, opposed to the assumption made by Eucken that the "electron" and "lattice" conductivities are additive. It is unfortunate that Peierl's theory does not permit of a comparison with experiment, and till this is possible it would be unwise to place too great weight upon it.

Of the wires used in this work the molybdenum and the monocrystal tungsten were kindly presented by the N.V. Philips Gloeilampen Fabrieken te Eindhoven (Holland) and the iron wire (with chemical analysis) by the Armco Corporation of America. A research scholarship by the University of Melbourne made it possible to carry out the work.

In conclusion it is a pleasure for the writer to express his thanks to Professor T. H. Laby, F.R.S., for the interest he has taken in this work, and for a number of suggestions.

---

* 'Z. phys. Chem.,' vol. 134, p. 220 (1928); vol. 125, p. 211 (1927); vol. 111, p. 431 (1924).

† 'Ann. Physik,' vol. 4, p. 121 (1930); vol. 5, p. 244 (1930).

Summary.

An electrical method of obtaining the thermal and the electrical conductivity of a metal in wire form, previously developed and used by the writer, is extended so as to include observations between −183° C. and 100° C. The wires studied are tungsten (in monocrystal form), molybdenum, silver, and iron. The complete thermal and electrical data obtained are given in Table II. Optical spectroscopic analyses of the purity of the tungsten, molybdenum, and silver wires were made. The nature of the impurities and, in the case of molybdenum and tungsten, their probable amounts are given in Table III of the Appendix.

APPENDIX.

By C. E. EDDY, D.Sc., F.Inst.P., and T. H. ODDIE, B.Sc.

Optical Spectroscopic Examination of the Wires.

A general qualitative analysis of a sample of the tungsten, molybdenum, and silver wires used was carried out, using a Hilger E2 quartz spectrograph. The spark spectrum was excited by connecting electrodes of wire in parallel with a condenser (capacity 0·03 μF) which was placed in series with an inductance (1 mH.) and the secondary of a transformer giving 15,000 volts.

A search of the spectra was made only for the "raies ultimes" and the stronger lines of the majority of the metallic and semi-metallic elements, using a range of wave-length 2200 A. to 3350 A. If the "raies ultimes" were not observed an element was regarded as not being present in an amount sufficient to be detected by a method employing spark spectra.

The "raies ultimes" of de Gramont,* and strong lines from Kayser's† tables were used in identification, together with lines recommended by Negresco‡ and by Lundegardh.§

The impurities found in molybdenum and tungsten wire are given in Table III.

Although no quantitative determinations of the amounts of the impurities present have been carried out, it has been possible to estimate approximately the amount of an element present by using the data which has been published

---

* 'C. R. Acad. Sci. Paris.,' vol. 166, p. 94 (1918); vol. 170, p. 1037 (1920); vol. 171, pp. 1106, 1129 (1920); vol. 173, p. 13 (1921); vol. 175, p. 1025 (1922); etc.

† "Tabelle der Hauptlinien der Linienspektra Allen Elemente" (1926).

‡ 'Thesis' (Paris, 1927).

§ "Die Quantitative Spektralanalyse der Elemente" (1929).

Table III.—The Impurities are expressed in per cent.

	Molybdenum.	Tungsten.		Molybdenum.	Tungsten.
Al	0·01	—	Pt	0·001	0·001
Bi	0·05	—	Rh	0·001	—
C	Trace	—	Sc	—	Trace
Cd	0·05	—	Si	—	0·01
Co	0·001	0·001	Sn	0·01	0·001
Cr	—	0·001	Ta	—	0·01
Cu	0·001	—	Ti	0·01	—
Ge	0·01	—	V	0·01	0·01
In	—	0·001	W	0·01	—
Os	—	0·001			

by various workers * relating the intensity and the first appearance of certain lines of each element with the amount of the element present. The percentages of each element, as given in the fourth column, are not claimed as having a high accuracy, but should be considered rather as the upper limit of the amount which is present.

The silver wire was drawn from a rod of "H.S." brand silver supplied by Messrs. A. Hilger, Ltd. Spectroscopic tests in this case were carried out simply to demonstrate that no appreciable impurities were introduced from the dies during the drawing of the wire. Spectra of the original silver and of the drawn wire were photographed together on the plate under equivalent conditions of exposure. Examination of the two spectra showed that no recognisable impurity had been introduced, so that the silver wire as used in the conductivity investigations was of the same purity as the original sample. The report of chemical and spectroscopic analyses supplied by Messrs. Hilger states that copper, lead, bismuth, magnesium, calcium, sodium, and silicon were the only elements present, and that these were present in an amount sufficient only to give one or two ultimate lines.

---

* Negresco, *loc. cit.*, and 'J Chem. Physique,' vol. 25, pp. 142, 216, 308, 343, and 363 (1928); Lundegardh, *loc. cit.*; Gerlach und Schweitzer, "Die Chemische Emission-spectral-analyse," Leopold Voss, Leipzig (1930); Lowe, "Atlas der letzen Linien der wichtigsten Elemente." Steinkopff, Dresden (1928); Smithells, "Tungsten," Chapman & Hall (1926); Smith, 'Trans Faraday Soc.,' vol. 26, p. 101 (1930); Ryde and Jenkins, "Notes on Ultimate Ray Powder," A. Hilger (1931).

The Deposition of Sputtered Films.

By R. W. Ditchburn, Professor of Experimental Philosophy, Trinity College, Dublin.

(Communicated by Sir Joseph Thomson, O M , F.R S —Received January 31, 1933.)

(1) *Introduction.*

The object of nearly all other work on cathode sputtering has been the investigation of the processes by which the metal is removed from the cathode.* The primary object of the present work is to investigate the deposition of the sputtered material on glass and metal surfaces placed near the cathode. It is known from the experiments of Knudsen† and other workers that a stream of atoms will not condense on a glass or metal surface unless the density of the stream (*i.e.*, the number of atoms striking 1 sq. cm. of the surface per second) exceeds a certain critical value. This critical density does not depend much on the nature of the surface unless the surface is entirely free from a gas layer, but it is a rapidly varying function of the temperature. The first of the present experiments was designed to see if these critical densities existed for metal sputtered on to insulated surfaces. No trace of this phenomenon was found, and further experiments were made to test different explanations of this result.

2. *General Notes on Experimental Conditions*

Before describing the conditions of individual experiments it may be convenient to give the following information relevant to all or most of the experiments.

(*a*) Cadmium was used as cathode in all experiments in order to be able to make a direct comparison with the detailed experiments of Cockcroft (*loc. cit.*) on the deposition of vaporized cadmium.

---

* Other observations on deposition have been made by workers interested in the production of mirrors and by Baum, 'Z. Physik,' vol. 40, p. 686 (1929). *Vide* also Blechsmidt, 'Ann. d. Physik,' vol. 81, p. 999 (1926).

† R. W. Wood, 'Phil. Mag.,' vol. 30, p. 300 (1915); Knudsen, 'Ann. d. Physik,' vol. 50, p. 472 (1916); Chariton and Semenoff, 'Z. Physik,' vol. 25, p. 287 (1924); Cockcroft, 'Proc. Roy. Soc.,' A, vol. 119, p. 293 (1928); Estermann, 'Z. Physik,' vol. 33, p. 340 (1925).

(*b*) In the above experiments there is no significant* difference between surfaces which have been roughly cleaned (*e.g.*, by emery) and those which have been outgassed before use. At the pressures needed for discharges it is impossible to keep surfaces free from gas, therefore, in most experiments the copper surfaces were merely cleaned with emery cloth or scraped with a razor blade before putting them into the vacuum. In a few experiments they were outgassed in a quartz tube before being put into the vacuum, but no difference in the results was obtained. A fresh piece of metal or glass was used to receive the deposit in each experiment.

(*c*) Precautions were taken to avoid grease. In some experiments the apparatus was arranged entirely without wax joints or ground joints, and was heated under vacuum to over 400° C before the experiment. It was necessary to let the vacuum down after this heating to insert the cadmium electrode. In the later experiments some joints with Apezion grease or Apezion wax were used.

(*d*) The cadmium used was stated by Messrs. Johnson Matthey (who kindly presented it) to be 99·9 per cent. pure. It was melted in a vacuum, and a little evaporated before use. The cathodes in general consisted of a mass of fairly large crystals.

(*e*) The gases used were hydrogen and argon. The hydrogen was admitted through a palladium tube, and in some experiments was circulated ; the argon was stated to be 99·5 per cent. pure, the chief impurity being nitrogen.

(*f*) Early experiments were done with ordinary self-maintained discharges. It was soon found to be necessary to introduce filament maintained discharges, in order to be able to work at low pressures. For this purpose a low voltage arc was used, positive ions from the arc being used for sputtering. After several types had been tried the arc shown in fig. 1, *a*, was found to be the most convenient. The cathode of the arc is a coated platinum strip filament (F) about 10 × 1 × 0·01 mm., and the anode is a nickel cylinder (K) (diameter about 3 cm., length 1 cm.), about six fine wires are fixed over the end nearest the main discharge. About 100 volts p.d. was used, and the current used was between 50 and 150 milliamperes. In general, the current was limited by external resistances so that it did not much exceed the emission of the filament. A little sputtering from the filament was observed, but it was always placed a long way from the cadmium cathode so that platinum could not be sputtered from it on to the cathode.

---

* The magnitudes of the critical densities are affected by the treatment of the surface, but not the general features of the phenomenon.

3. *The Deposition of Sputtered Material on Insulated Surfaces.*

(*a*) *The Absence of a Critical Density Phenomenon.*—Several preliminary experiments on slow sputtering were made. Different kinds of discharge tubes and bulbs were used, and pressures ranging from 10^{-3} to 10^{-1} mm.

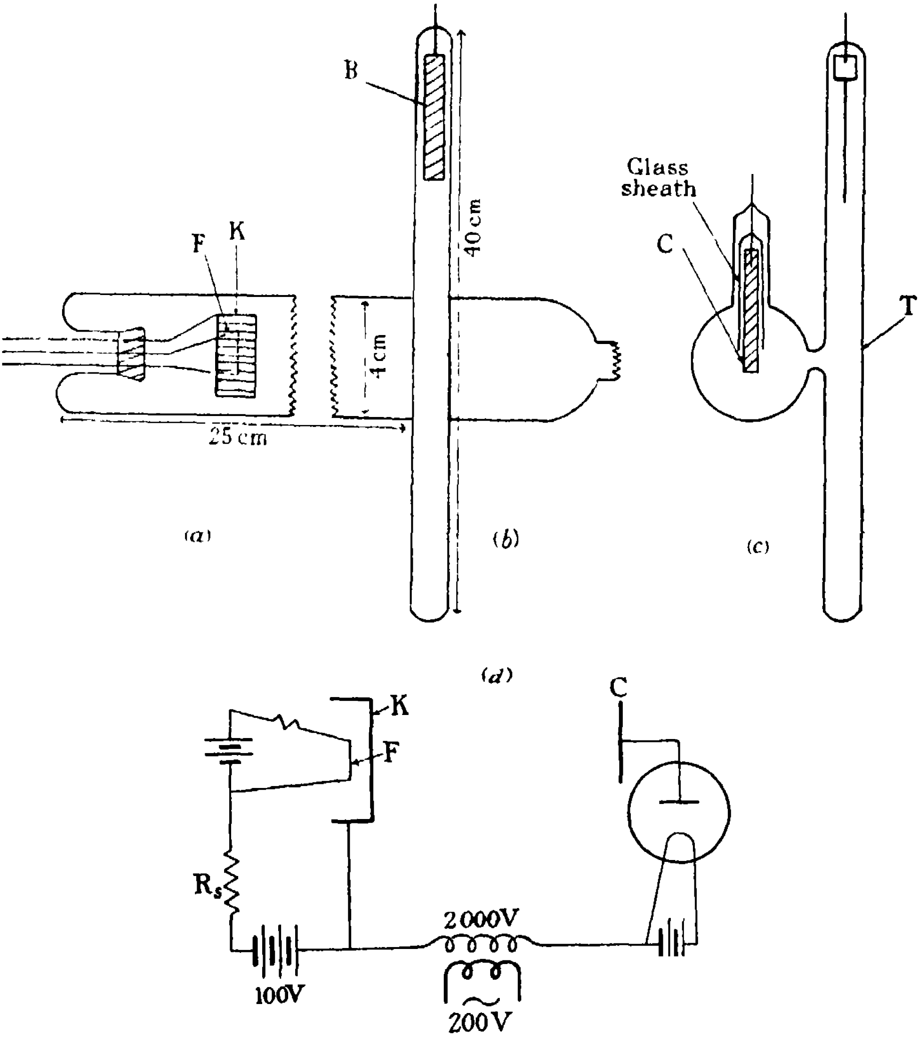


Fig. 1.—(Target system not shown)

The anode was an aluminium rod. The high tension was first supplied by a coil, but after one or two experiments a 3000 volts transformer was used. The current was rectified but not smoothed. This removes the possibility of metal being sputtered from the aluminium electrode on to the cadmium and thence to the walls. With currents of order 1 m.A. to 0·1 m.A. the sputtered film became visible in times varying from 1 minute to 1 hour. In some experiments films were deposited on glass, in others on insulated copper surfaces.

A film of cadmium deposited by the boiling process is just visible when 4 molecules thick.* For reasons, which will be given later, the writer believes that sputtered films do not become visible till about 5 or possibly 10 times thicker. Taking the highest estimate of the thickness and assuming a uniform rate of deposition the stream density is 10^{14} atoms/c.c.^{2}/sec. for films which take an hour to deposit. According to Cockcroft a stream of vaporized cadmium of this density would fail to condense on a glass or metal surface at any temperature above —140° C. The highest temperature of the receiving surface in Cockcroft's experiments was —90° C., and the critical density determined for this temperature was nearly 10^{16} atoms/c.c.^{2}/sec. It is not justifiable to extrapolate Cockcroft's curve and hence derive the critical density for room temperature, but this density should certainly be much greater than 10^{16} atoms/c.c.^{2}/sec.

Thus these experiments show that there is no critical density for deposition of sputtered cadmium of anything like the same order as the critical density for boiled films. This result was later confirmed by experiments in which hydrogen and argon were used, and in which a high tension battery of 1000 volts supplied the potential difference.

(*b*) *Rate of Deposition as a Function of the Temperature.*—The apparatus shown in fig. 1 was designed to test whether the rate of deposition is affected by the temperature of the target. In order to eliminate fluctuations in the discharge conditions it was arranged to have two targets in a side tube, and to bring them alternately opposite the cathode by means of a magnetic control. Ions from the low voltage arc described above (p. 170) are accelerated on to the cadmium cathode C by a potential of about 2000 volts (rectified but not smoothed). Two copper targets are mounted on two similar iron rods, the iron being recessed slightly to give room for the thin copper sheet. Down the centre of the rods passes a thin pyrex tube, and by means of tungsten wires sealed into the pyrex, it is arranged that the system is rigid, both relative rotation and translation being prevented. The whole system slides vertically in a tube T. A blade (B) fixed to the tube slides in a slot in the upper rod, thus preventing the system from rotating relative to the tube while permitting vertical translation. The two targets are thus heat insulated from one another, but firmly connected mechanically.*

---

* Estermann, 'Z. phys. Chem.,' vol. 106, p. 399 (1923).

The lower half of the tube T could be surrounded by liquid air. Before commencing an experiment liquid air was left on this tube overnight, and the target system was left in the down position, so that the lower rod was almost completely surrounded by liquid air. Allowing only for cooling by radiation this should give ample time for cooling to below —140° C.

The experiment was first performed with hydrogen at a pressure of 7×10^{-3} mm. Alternate 2-minute exposures were given. A deposit appeared on the surface at room temperature after a total exposure of 45 minutes, and on the cooled surface after a total exposure of 50 minutes. In view of the fluctuations in discharge conditions the difference in time is not significant.

The experiments with hydrogen were not regarded as entirely satisfactory. The discharge caused a very rapid disappearance of gas, and the hydrogen caused large and irregular changes in the emission of the filament. Also it seemed just possible that some cadmium hydride might be formed, and that this might assist deposition. The experiment was therefore repeated using argon. The disappearance of gas was slower, and conditions were much more steady. It was now found that the deposit became visible on the cooled surface first. The ratio of the time for the appearance of a deposit on a surface at —140° C. to that for the appearance of a deposit on a surface at +20° C. was about 1 : 1·3 (a difference of 30 per cent.).

This appeared to indicate that deposition was definitely more rapid on the cooled surface, but it was observed that the deposit on the cooled surface seemed to disappear when the surface was heated to room temperature. It again became visible (though less clearly) on re-cooling.† This means that the deposit on the cooled surface is different in some property (*e.g.*, structure) which renders it more easily visible than that on the warmed surface. If any difference in the rate of deposition exists it is certainly not greater than 30 or 40 per cent. for a temperature difference of 160° C. For vaporized metal a difference of 20° C. alters the critical density by a factor of about 10. Some experiments were tried with surfaces heated above room temperature to see whether deposition could be prevented. These did not give any definite result. When the temperature is over 200° C. deposition apparently becomes slower simply owing to thermal re-evaporation. Rough tests showed that the rate of deposition on glass is not much affected by large changes of temperature.

---

* Other methods of connection depending on screws failed owing to slips caused by repeated temperature changes.

† This effect is possibly due to a condensed film of argon on the cooled surface. It was not noticed with hydrogen.

(*c*) *Control Experiments with Vaporized Cadmium.*—The experiments of Cockcroft and others were all done with vapour in high vacuum. It appeared possible that the difference between the behaviour of boiled and sputtered metal might be caused by the presence of the gas, and not by the existence of the discharge. It was therefore decided to do experiments with vaporized cadmium with a gas present at pressures as high as those employed in the experiments described above. The presence of gas rendered it difficult to calculate the stream density; but there was no doubt of the existence of the critical density phenomenon. In one experiment a deposit appeared on the cooled surface in 2 minutes, the temperature of the cadmium being 260° C. The surface at room temperature was exposed for 30 minutes, the temperature of the cadmium being raised gradually from 260° to 315° C. No deposit appeared although the stream density at 315° C. must be many times greater than at 260° C.

4. *Experiments on the Condensation of Vaporized Metal in Discharge Tubes.*

The above results appear to be due to one of the following causes (or to a combination of them):—

(1) The particles sputtered from the cathode may differ in some way (*e.g.*, charge, mass, velocity) from those produced by vaporization.

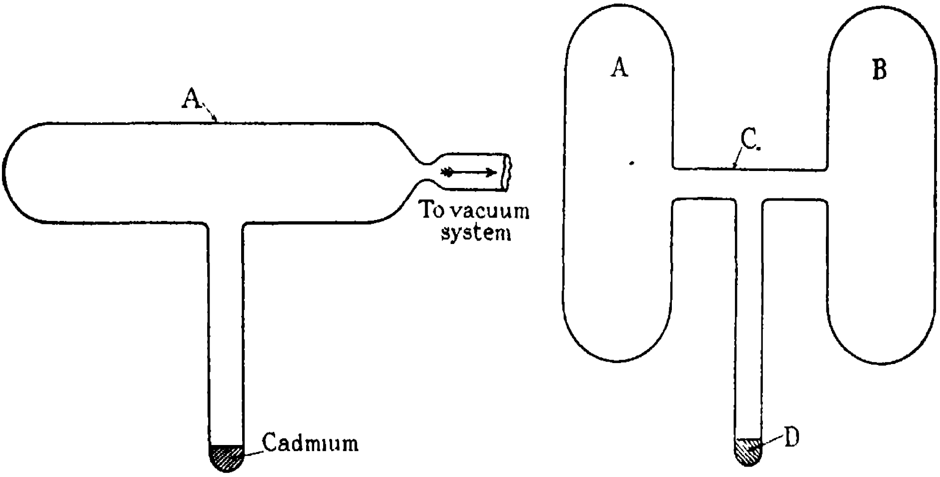
(2) The particles may be altered (*e g.*, by being ionized) during their passage from the discharge tube to the wall.

(3) The discharge may affect the surface in some way so as to assist the condensation of the cadmium.

If the absence of the critical density phenomenon is caused entirely by (1) we should expect that on passing cadmium vapour into a discharge-tube it would fail to condense unless the stream density were greater than the critical density. If it arises partly or wholly from (2) or (3) we should expect the vapour to form a deposit in a discharge tube under the conditions of formation of sputtered deposits. The following experiments were therefore carried out.

(*a*) A discharge tube with low voltage arc, similar to that shown in fig. 1, was constructed. The cathode C was replaced by a copper rod to serve as target, and the double-target system was omitted. The rod was left insulated. Cadmium was heated in a side tube; the heating being started some time before the discharge. No condensation took place on the target, though metal condensed on parts of the wall which had been sensitized by touching with liquid air—thus showing that there was sufficient vapour present to form a deposit. On switching on the low voltage arc discharge a deposit was quickly formed on the copper and on part of the wall not previously sensitized.

(*b*) A discharge tube of the simple form, shown in fig. 2*a*, was constructed. Cadmium was boiled in a side arm. The gas pressure was about 10^{-3} mm. The furnace was heated for half an hour without the discharge—no deposit being formed. An electrodeless discharge produced by high frequency oscillator (10^8 cycles/sec.) was then started in the main tube. A deposit appeared in 10 minutes at A. It failed to spread any further in an hour.



a. *b*

FIG. 2

(*c*) A double tube of the form, shown in fig. 2*b*, was constructed, a small piece of cadmium being placed at D, the tube evacuated to below 10^{-4} mm., and sealed off. On heating D and running the electrodeless discharge in A, the deposit first appeared at C, then when C was heated the deposit moved into A. It was then distilled backwards and forwards in A for some time, no deposit appearing in B. In this experiment the part of the tube where there is no discharge is always receiving as many cadmium atoms as that on which the deposit actually occurs. These experiments show that the discharge, both under the conditions of the earlier experiments and at very low pressures, is able to cause the condensation of the boiled metal on an insulated surface.

It also appears that under these conditions every cadmium atom condenses at the point where it first strikes the wall (provided that point is within the region of the discharge).

(*d*) *Test for the Presence of Charged Particles.*—The work of Baum,* etc., has shown that sputtered particles leave the cathode uncharged. Under the conditions of Baum's experiments a considerable proportion became charged a few millimetres from the cathode. It is not likely that many particles would become charged under the conditions of the present experiments. The current density in the main discharge is very small (10^{-4}–10^{-5} amps./cm.^{2}) most of it is carried by positive ions of the gas, and electrons from the low voltage arc cannot penetrate into the region near the cathode. Since, however, a small proportion of charged particles might stick to the surface and act as condensation centres it appeared desirable to check the point by direct experiment. Two experiments were carried out In the first the sputtered particles were made to pass along a tube of about 1 cm. radius through a magnetic field of about 800 gauss extending over 6 cm. of their path. According to Baum the charged particles are atoms of energy 0·1 volt, and such particles should be bent into a circle of a few millimetres radius by this field. Thus nearly all would hit sides of the tube, and fail to pass the magnetic field. Definite deposits were formed beyond the magnetic field. It was not possible to determine whether they appeared as quickly with the field on as with it off because the stray field disturbed the discharge.

In a second experiment the particles had to pass through an electric field produced by two plates 4 mm. apart with 40 volts p.d. across them. The mean potential was that of the cathode. The length of path in the field was about 5 mm. Thus the slow charged particles should be drawn into one or other of the plates, according to the sign of their charge. A piece of glass was placed so that part of it received particles which had passed through the field, and part received particles which had passed near the plates, but not through the field. Deposition on the two parts of the glass was equally rapid. Thus these experiments show that the absence of critical densities cannot be explained by the presence of charged particles unless these particles are sufficiently fast to pass through our fields.†

5. *Deposition on Surfaces at Different Potentials.*

In view of the results of the previous experiments it appeared that the most probable cause of the absence of critical densities lay in some action of the discharge on the surface. Presumably such action would be due to positive ion or electron bombardment. In order to test this theory a tube containing a low voltage arc and cadmium cathode, as in fig. 1, without the side tube, was contructed. A copper wire connected to the anode was used as target. It was found that the sputtered material did not condense on this wire when it was at room temperature. In order to test whether condensation of sputtered material on such a surface could be produced by cooling, a copper rod was placed in a side tube, and cooled overnight. It was connected to the anode by a fine wire. A small piece of copper-foil was placed between the cathode and the copper rod, shielding part of the latter. This was electrically connected to the anode, but not cooled. A deposit appeared on the cooled rod in about 5 minutes. This deposit had a fairly sharp edge showing the shadow of the foil. No deposit appeared on the foil for about 3 hours when a very faint coloured deposit became visible. The existence of the shadow and its sharpness demonstrates clearly that the sputtered particles were striking the foil and not being condensed. It also shows the pressure was low enough to make collisions with gas molecules rare.

---

* 'Z. Physik,' vol. 40, p. 686 (1929).

† The above result applies, of course, only to experiments at low pressures with the low-voltage arc. It is not in conflict with Baum's results.

In order to investigate this effect further a series of experiments on both vaporized and sputtered metal were carried out. Fig. 3A shows the apparatus used for the experiments on vaporized metal. Ions from the low voltage arc are drawn out of the tube A into the main bulb, and accelerated on to the cathode C. This consists of an aluminium plate with a hole in the centre. A rough beam of vapour from the quartz tube Q is fired through this hole into the centre of the main bulb. The quartz tube is supported and heated by a helix of

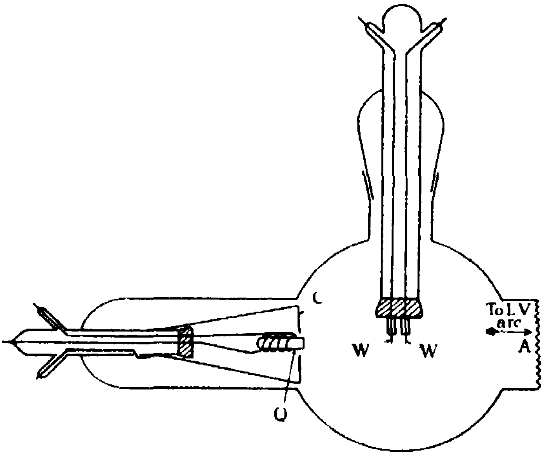


FIG. 3A.—Ground joint to be rotated through 90° when in use.

tungsten strip. The mouth of the furnace is just in front of the cathode so that little cadmium reaches it directly. The deposition of the metal on fine copper probe wires (W) was observed. Five probes were used (for clearness of reproduction two only are shown in the diagram). Each probe consists of a copper wire about 5 mm. in diameter projecting from a pyrex tube which shields it to within about 5 mm. of the end. For the experiments on sputtered cadmium the aluminium cathode and the quartz tube were replaced by a cadmium plate (made by moulding the cadmium in vacuum) of similar size. The position of this plate could be adjusted by a magnetic control.

The electrical connections are shown in fig. 3B. The main potential is supplied by a high tension battery. The resistances R_s are for safeguarding the batteries. The current to the cathode C and the current in the low voltage arc, are measured by two milliametres (m.A.). The probes (P_1, etc.) are each connected through a resistance (R_1, etc.) to a subsidiary battery. By inserting the plug Pt into each of the sockets (S_1, etc.) in turn, it is possible to read the current to the different probes in turn. By suitably choosing the values of the shunting resistances R_1, R_2, etc., currents of very different order may be read on the same galvanometer (G). Before commencing the main experiments preliminary electrical measurements were made to determine the current-voltage characteristic for one of the probes under the discharge conditions which

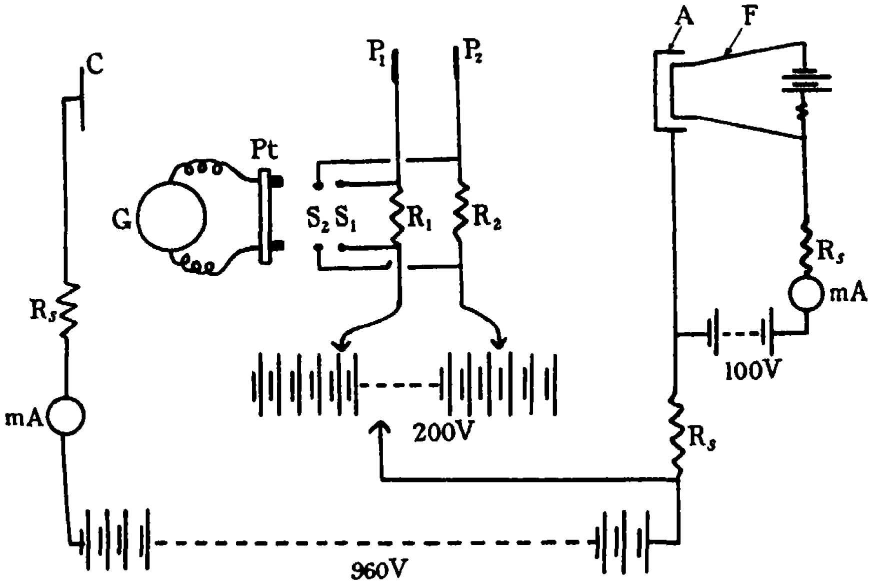


Fig. 3B.

were to be used in the main experiment. The variation in the current collected by one probe produced by altering the potentials of its neighbours was also investigated. In general, this last effect was found to be fairly small. As a result of these measurements the potentials to be used in the main experiment were chosen. The potential at which the probes collected no current was determined. Measuring from this potential as zero, one probe was placed at a negative potential of over 100 volts, one at a positive potential of 30 volts or more, two others were placed at intermediate potentials, and the fifth probe was left insulated. The currents to the different probes and the pressure were measured every 3 or 4 minutes during an experiment. Experiments were done at different current densities and different pressures, both with sputtered and vaporized metal. No difference between the behaviour of the two under strictly comparable conditions was observed. It was necessary to adjust the gas pressure to within a moderately narrow range—especially with the boiled metal—to avoid subsidiary effects. If the pressure became too high ($> 10^{-2}$ mm.) the effusion of the boiled metal ceased to be directed at all, and it was necessary to increase the temperature of the furnace to compensate this. This meant that the cadmium atoms reached the target by devious paths and had a considerable chance of becoming ionized. At low pressures ($> 2 \times 10^{-4}$) the discharge became very feeble; there were few positive ions present, and electrons from the low voltage arc penetrated into the region near the mouth of the furnace. (The penetration of beams of electrons could be seen visually). It is probable that under these conditions the number of cadmium atoms was comparable with the number of gas molecules, so that an appreciable fraction of the current was carried by positively or negatively-charged cadmium. Thus at pressures outside the range 10^{-2} mm. to 2×10^{-4} mm. the cadmium went to the most strongly positive and also the most strongly negative electrodes, much less being deposited on electrodes at or near the potential corresponding to no current. (Various minor observations support this explanation of the experiments at very high and very low pressures.)

Most of the experiments were done at pressures round about 1×10^{-3} mm. to 5×10^{-3} mm. For all these experiments, whether boiling or sputtering, the deposits appear in the following order:—

(1) Surfaces collecting slow positives (up to 30 volts).

(2) Surfaces collecting little or no current, and therefore receiving approximately equal number of positives and negatives.

(3) Surfaces collecting faster positives (order of 100 volts).

(4) Surfaces collecting large electron currents.

In all experiments the deposits received on surfaces collecting electrons were very small, and in some they could not be seen at all. It was noted that these differences between surfaces at different potentials were not obtained with sputtered platinum.

It has been suggested to the writer that the similarity between the effects with vaporized and sputtered particles in a discharge tube arises from the following indirect process. Some of the particles vaporized into the tube may condense on the cathode and be sputtered, serving as nuclei for condensation of other vaporized particles. While this process might just possibly explain the results described in this section, it could not account for those of section 4 (*b*) and 4 (*c*), in which an electrodeless discharge was used. For this and other reasons it does not appear probable that this indirect process is of much importance.

6. *Structure and Thickness of the Deposits.*

The structure of sputtered films has been subjected to X-ray analysis, and has been the subject of some controversy. The general conclusion* for films of platinum, gold, etc., is that the films are amorphous if formed on a cold surface, though they become crystalline if heated, and may sometimes be heated sufficiently by the discharge to make them crystalline. These results cannot, with certainty, be applied to cadmium. Because of the great difference in the critical density the process of condensation of sputtered gold, etc., may be essentially different from that of sputtered cadmium. Cockcroft (*loc. cit.*) has shown that the films of cadmium produced by slow boiling consist of small crystals showing the Maxwell-Garnet resonance colours.† The resonance colours may be distinguished from interference colours by immersing the surface in a liquid. Owing to the change in the refractive index of the medium the resonance colour changes, the effective wave-length increasing. Only on one occasion were these colours definitely found. This was a deposit formed on a slightly cooled surface which had been fixed at such a potential as to repel positive ions. It will be seen later that we believe such deposits are formed in a different way from ordinary sputtered deposits on insulated surfaces. The coloured deposits often seen in discharge tubes appear on test to be interference colours.

---

* Debinska, 'Bul. Pol. Acad. Cracow,' vol. 9–10A, p. 460 (1930); 'Z. Physik,' vol. 54, p. 46 (1929); Cau, 'Ann. Physique,' vol. 11, p. 354 (1929); Ingersoll and Hanawalt, 'Phys. Rev.,' vol. 34, p. 972 (1929); Dreisch and Rutten, 'Z. Physik,' vol. 60, p. 69 (1930).

† 'Phil. Trans.,' A, vol. 203, p. 385 (1904). See also R. W. Wood, "Physical Optics," p. 630 (2nd edn.).

Another method of detecting the small crystals is by the thickness of the film when just visible. A non-homogeneous film (of the type produced by boiling) became visible when the average number of molecules per unit area is only sufficient to provide a layer 3 or 4 molecules deep.* A homogeneous metallic film† does not acquire sufficient reflecting and absorbing power to make it clearly visible until it is about 20 or 30 molecules thick. The thickness of sputtered films when clearly visible‡ was determined by weighing. The films for this purpose were deposited on quartz coverslips (to avoid hygroscopic effects). Tests were made to see that there was no increase in weight if the quartz was placed in a discharge tube with an aluminium cathode (giving practically no sputtering). These control experiments were all satisfactory. The weight was 3×10^{-5} gm. per cm.^{2}, and assuming that the density of the film is about the same as the density of cadmium in bulk the depth is $3 \cdot 7 \times 10^{-6}$ cm., *i.e.*, about 90 molecular diameters. This taken with the absence of resonance colours indicates a homogeneous and probably amorphous film in agreement with X-ray analysis of films of other metals.

7. *Other Observations.*

The following incidental observations were made in the course of these experiments.

(*a*) *Penetration of Sputtered Cadmium into the Glass.*—Two pieces of evidence show that some sputtered material penetrates into the glass wall of the discharge tube. Firstly, even after cleaning deposits off thoroughly with hot chromic acid, it was often very difficult to seal a portion of the tube to fresh glass.§ Secondly, on one occasion a tube in which cadmium had been sputtered was cleaned with chromic acid, and heated for some time to 400° C., in air; on subsequently passing cadmium vapour into the tube it condensed first on parts on which the cadmium had previously been sputtered—indicating a persistent sensitization of the surface.

(*b*) *Attempts to Sputter at Low Pressures.*—Two attempts were made to do sputtering experiments at very low pressures. In the first of these K^+ ions from a Kunsman filament were accelerated by 2000 volts on to a cadmium cathode. Only small currents (order of 0·01 milliamps.) were available, but the experiments were continued for over 10 hours. It was calculated that with argon, currents of this size should produce detectable sputtering in 2 hours. No sputtered deposit was obtained, but after about 1 hour a faint blue deposit appeared on the cadmium surface. The number of K^+ ions received during this time was sufficient to form a layer 3 or 4 molecules deep. Thus under these condtions the K^+ ions appear to stick to the cathode, instead of removing cadmium from it.

---

* Estermann, 'Z. phys. Chem.,' vol. 106, p. 399 (1923).

† R. W. Wood (*loc. cit.*), p. 467.

‡ There was great difficulty in deciding when a film first became visible. Watching for the appearance of a film one would suspect a deposit for some time without being quite sure of its existence. Then suddenly the film would be seen to be quite strong. At this stage the discharge was stopped, and the film weighed.

§ Dr. Emeléus tells me he has also observed this effect.

In the second experiment a small discharge tube with cadmium electrodes was pumped out and sealed from the pump. No discharge could be obtained with an induction coil, but one was produced by a high frequency oscillator. A sputtered film was produced in this way in about 15 minutes. This would appear to be the most convenient way of sputtering at very low pressures.

(c) *Surface Motion of Sputtered Particles.*—The surface motion of particles which have been condensed from cadmium vapour was demonstrated by Estermann[*] and Cockcroft (*loc. cit.*). By an experiment, similar to that of Estermann, the writer[†] has obtained strong, though not conclusive, evidence that surface motion exists for cadmium sputtered on to glass.

8. *General Discussion.*

The Adsorption Process.—The experiments of Cockcroft and others on the condensation of boiled metal were adequately explained by the theory of Langmuir[‡] and Frenkel.[§] According to their view, the surface forces of ordinary glass, and metal surfaces are partially neutralized by the presence of a thin film of gas or vapour. A cadmium atom striking such a surface is not immediately adsorbed, but moves about on the surface like a molecule in a two-dimensional gas. After a certain average time (τ), the atom leaves the surface, unless it previously meets another cadmium atom, and combines with it, when the pair will have a much longer life on the surface. From these assumptions it is shown that critical densities must exist, and the theory also explains the variation of the critical density with temperature. The critical density phenomenon should not exist on a perfectly gas-free surface, for with such a surface the cadmium atoms would be "captured" by the surface

---

ann, 'Z. Physik,' vol. 33, p. 320 (1925).

[†] Ditchburn, 'Proc. Camb. Phil. Soc.,' vol. 29, p. 131 (1933).

[‡] Langmuir, 'J. Amer. Chem. Soc.,' vol. 40, p. 1361 (1918).

[§] Frenkel. 'Z. Physik,' vol. 26, p. 117 (1924) (quantitative theory).

forces as they arrived, and a continuous film built up. Cockcroft was not able to produce a perfectly gas-free surface, but he was able to show that the critical density is several times smaller on a recently formed surface of silver than on surfaces which had been degassed and then cooled* (thus allowing them to acquire a layer of gas).

Since we do not find critical densities under ordinary conditions with sputtered particles the process of adsorption must in some way be modified by the discharge. The explanation first considered by the writer was that cadmium produced by sputtering contained a fairly small proportion of diatomic molecules, or larger aggregates, or, alternatively, that it contained some fast-moving atoms. The aggregates would, obviously, form nuclei of condensation. The fast moving atoms might penetrate the gas layer and be captured, and then act as nuclei. This theory had to be abandoned when it was found that the condensation of sputtered particles is identical with that of cadmium vaporized in a discharge tube.

The next theory is that the discharge may alter the vapour by producing excited or ionized atoms. We have shown directly that there are very few charged particles present of the type found by Baum, and that the condensation is not dependent on their presence. We shall consider the question of faster charged particles later. Electrons possessing more energy than that required to ionize the cadmium atoms would produce more ions than excited atoms. From the general theory of the discharge it is known that most of the electrons present would have more than this energy. Thus since we have failed to detect ionized atoms, there cannot be any appreciable number of excited atoms present. Also any theory based on the action of excited atoms must fail to explain the difference between condensation on surfaces at different potentials.

We are, therefore, left with the idea that the discharge may affect the surface so as to assist the adsorption. Positive ions, electrons and fast-moving atoms —produced by recombination—are all able to heat the surface. The electrons will penetrate into the metal, and their heating effect will thus be distributed, on the other hand, the positive ions, may be expected to produce a strong local heating on the surface. Miss Blodgett and others have shown† that bombardment by fairly slow positive ions produces a considerable evolution of gas from the walls and electrodes of discharge tubes. Thus we may explain our results in the following way. In a discharge tube, surfaces which are receiving positive ions are continually being cleaned; this cleaning effect is counteracted by a fairly rapid condensation of molecules from the gas present in the tube. Thus at any one time, a certain fraction of the area of the surface will be free from gas. When a cadmium atom strikes such a surface, it will move about until either (*a*) it reaches one of the cleaned areas, when it will be captured by the surface forces, or (*b*) it meets another cadmium atom, and forms a pair, or (*c*) if neither (*a*) nor (*b*) occur within an average time (τ) it will leave the surface. When we are dealing with conditions under which the rate of arrival of the atoms is far below the critical rate (*b*) may be neglected. The deposit will then be built up by (*a*) alone. It is obvious that if there are sufficiently large numbers of positive ions striking the surface, the film will be built up, even though the density of the stream of atoms striking the surface may be far below the critical density for deposition in the absence of the discharge. Under most discharge conditions insulated surfaces receive considerable numbers of positive ions, and are therefore in a condition to receive deposits. If, however, a surface is made a few volts positive to the space, positive ions will be repelled, and the cleaning effect of the electrons and neutralized ions is apparently too small to be effective. But if the temperature be reduced sufficiently, a film of sputtered cadmium may be built up on such a surface, mainly by process (*b*). If the surface be made strongly negative to the space, then large numbers of positive ions will be drawn in, but with a high velocity. Cadmium atoms condensing on such a surface may be immediately re-sputtered, so that a deposit cannot be built up on such a surface so rapidly as on one which receives slow positive ions.

---

* This point is also indirectly confirmed by experiments on thermionic properties at high temperatures where gas-free surfaces can be produced and maintained.

† 'Phil. Mag.,' vol. 4, p. 168 (1927).

Thus we expect that both for boiled and sputtered metal, the deposits will appear, first, on a surface slightly negative to the space (and therefore receiving large numbers of slow positive ions), second, on an insulated surface (receiving a smaller number of positive ions), third, on a strongly negative surface (receiving fast positive ions), and last on a surface positive to the space (receiving electrons). This is the order found experimentally (see p. 179). According to this view we expect the film built up on the surfaces receiving positive ions to be continuous, because the deposit is not being built up by the primary formation of nuclei. (The cleaned areas on the surface are indifferently distributed and are continually changing.) This agrees with results of experiments described in Section 6. The existence of surface motion is also explained, Section 7 (*c*). The penetration of the surface by sputtered particles, Section 7 (*b*) is probably caused by a certain number of atoms being struck by positive ions while they are on the surface, and driven into it. As stated above (p. 177), films apparently of a different character are very slowly built up on surfaces which do not receive positive ions. Such films may be due to a small cleaning effect of the electrons,* or to the existence of a very few particles capable of acting as nuclei. If the rate of arrival of the sputtered atoms is above the critical density, we should expect that the deposition would be unaffected by the potential of the target. This condition may be realized for platinum at room temperatures, even for slow rates of sputtering. To attain the corresponding condition for cadmium, we must cool the surface to liquid air temperatures. With both metals the expected result was observed (pp. 180).

The above discussion is entirely qualitative. It is difficult to make any quantitative check on the theory, but the following considerations appear to show that it is not unreasonable from a quantitative point of view. With a current of density about 10^{-4} amps./cm.^{2}, the number of positive ions striking the surface per square centimetre per second is about 10^{15}. At a pressure of 10^{-3} mm., the number of gas molecules per square centimetre per second striking the surface is about 5×10^{17}. Thus if each positive ion cleans an area of atomic dimensions, and every gas molecule which hits a clean part of the surface was captured, we should expect that a fraction of about 1 in 10^3 of the surface to be clean when equilibrium had been reached. The experiments of Cockcroft show that the life of a single atom on the surface is not less than $2{\cdot}6 \times 10^{-9}$ seconds. Other considerations based on the Frenkel theory indicate that it is probably not more than ten times greater. A cadmium atom moving with thermal velocity corresponding to room temperature would pass over an area of about 10^{-12} sq. cm. in $2{\cdot}6 \times 10^{-9}$ seconds (*i.e.*, over about 10^3 atoms). It would thus stand a fair chance of being captured. We should expect that if the positive-ion current density were reduced much below 10^{-5} amps. per square centimetre the rate of deposition should be considerably reduced. This was observed.

In the above we have assumed that the faster positive ions were able to remove the cadmium by re-sputtering it, and so to reduce the effective rate of condensation. It is not possible to decide whether the observed currents were large enough to produce such effects because there are no data on the sputtering of cadmium at voltages under 1000 volts. An attempt was made to remove one of the deposited films by bombardment with slower positive ions. An electrode on which a cadmium deposit had previously been formed was set at

---

* Some cleaning effect might also be produced by negatively charged gas molecules, or by fast neutral molecules formed by the recombination of ions in the discharge.

100 volts negative to the space, and then the discharge was adjusted so that a current of 3×10^{-5} amps. was collected, but no disintegration or disappearance of the deposit was observed in 2 hours. This observation offers some evidence against the view that the effective rate of deposition can be reduced by positive ion bombardment. This evidence is not, however, very strong because more energy is required to remove a cadmium atom from a continuous cadmium surface than to remove a " free " cadmium atom which is still moving about on a copper or glass surface.

Two alternative theories should be mentioned.

(*a*) It is possible that the condensation caused by positive ions is due to a change in the electrical state of the surface produced by the ion as it enters the surface. But the time during which a positive ion is near to the surface (*i.e.*, within a few atomic diameters) is so exceedingly small that it cannot give an explanation of the magnitude of our results unless we assume that the electrical change persists for some time after the ion has struck the surface. Since there is no other evidence for any such effect it appears very improbable that this view can be correct.

(*b*) While the experiments of section (4) show that the deposition is not affected by slow cadmium ions such as those found by Baum (*loc. cit.*) they do not exclude the possibility of a sensitizing effect caused by cadmium ions of velocities corresponding to 5 volts or more. It is very difficult to design any experiment which will clearly discriminate between an effect of these ions and one due to A^+ ions from the discharge, since any potential which will repel one will also repel the other. We have, however, no positive evidence for the existence of these faster cadmium positive ions, and it is certain that their number must be very small compared with the number of A^+ ions present, and also the number of cadmium atoms present.

The structure of the deposits also offers fairly strong evidence against any theory based on Cd^+ ions, or on any other local sensitizer. If the formation of the deposit depended on the presence of nuclei we should expect it to be formed round these nuclei, *i.e.*, we should expect deposits consisting of small aggregates similar to those found by Cockcroft. As stated in Section 6 the deposits found are continuous, and therefore not formed by means of nuclei. Thus the theory we have given above (p. 183) is supported by considerable evidence, and would appear to offer the best working hypothesis for more detailed experiments.

B.—*The Sputtering Process.*—Although the primary object of these experiments was not the investigation of the sputtering process, it is of interest to consider whether the results have any bearing on this problem. It will be seen that the behaviour of the sputtered particles is in every way similar to that of particles boiled into the discharge tube. This confirms the conclusions of other workers* that the sputtered particles leaving the cathode are uncharged single atoms with velocities of the order of the thermal velocities for under 1000° C. If there were present more than a few per cent. of molecules or larger aggregates, we should expect them to form nuclei of condensation, when the behaviour of the sputtered particles and the boiled particles would not be so closely alike. In particular, the sputtered particles might be expected to condense much more rapidly on the surfaces which receive electrons. The experiments of von Hippel and Baum would probably fail to reveal the existence of 20 per cent. molecules, but a much smaller number should have been revealed by our condensation experiments.

Our results are thus consistent with the "local boiling" theory of von Hippel, but they cannot be regarded as supporting that theory, since other mechanisms of sputtering might equally well lead to the production of single uncharged atoms with thermal velocities.

I am glad to take this opportunity of thanking Sir J. J Thomson for his continued interest in my work. I wish also to thank Dr. Braddick and Dr. K. G. Emeléus for helpful discussions while the work was in progress.

Summary.

(1) It is shown that in the deposition of sputtered cadmium there is no critical density phenomenon similar to that found by various authors for the condensation of vapour. For insulated surfaces the rate of deposition is very little affected by temperature changes from +30° C. (or more) to −150° C.

(2) The rate of deposition is affected by the conditions of the discharge and by the potential of the target, being faster for surfaces which are collecting slow positive ions (order of 10 volts), somewhat slower for those collecting fast positive ions (order of 100 volts), and almost indefinitely slow for surfaces collecting electrons.

(3) The rate of deposition on surfaces collecting electrons can be greatly accelerated by cooling.

---

* von Hippel, 'Ann. Physik,' vol. 80, p. 672 (1926); vol. 81, p. 1043 (1926); Blechsmidt, 'Ann. Physik,' vol. 81, p. 999 (1926); Baum, 'Z. Physik,' vol. 40, p. 765 (1929).

(4) It is shown that the deposition of boiled cadmium can be affected very strongly by a feeble discharge, and effects exactly parallel to those mentioned in (2) are found.

(5) The sputtered metal giving these effects is shown to be uncharged.

(6) Experiments on the thickness and structure of the film are described.

(7) A tentative theory is suggested according to which the deposition is assisted by a surface cleaning effect of positive ions.

Determinations of the Signs of the Fourier Terms in Complete Crystal Structure Analysis.

By K. Banerjee, D.Sc.

(Communicated by Sir William Bragg, O.M., F.R.S.—Received February 1, 1933.)

The Fourier representation of the results of X-ray diffraction by crystals was first suggested by W. H. Bragg,* and later it was used by Havighurst and Compton† to obtain the distribution of scattering matter in crystals. W. L. Bragg‡ developed the two-dimensional representation of the Fourier series into a very convenient method of crystal analysis. He considered crystals having a centre of symmetry where it can be shown that if a projection of the atoms on the yz plane be considered the density of scattering matter per unit area of projection at the point y, z is given by

$$\rho\,(y, z) = \frac{1}{bc \sin \alpha} \sum_{-\alpha}^{+\alpha} \sum_{-\alpha}^{+\alpha} \mathrm{F}\,(0kl) \cos 2\pi \left(\frac{ky}{b} + \frac{lz}{c}\right) \qquad (1)$$

the structure factor for X-ray diffraction from the plane $(0kl)$ of the crystal being the numerical value of $\mathrm{F}(0kl)$.

The sign of $\mathrm{F}(0kl)$ may be either plus or minus, but cannot be determined from diffraction intensities. The method that is being used by W. L. Bragg and his students is to determine the structure very roughly by the trial and error method with guidance, if available, from physical and chemical properties. The signs of the terms are then determined from this rough structure and the complete analysis effected by the Fourier method, with the help of these signs with, if necessary, slight modifications.

---

* 'Phil. Trans.,' A, vol. 215, p. 253 (1915).

† Havighurst, 'Proc. Nat. Acad. Washington,' vol. 11, p. 502 (1925); Compton, "X-rays and Electrons," p. 151.

‡ 'Proc. Roy. Soc.,' A, vol. 123, p. 537 (1929).

Bradley's* important discovery of the variations of scattering power for different wave-lengths has made it possible to determine the signs beforehand for simple crystals composed of two fairly heavy elements.

All the methods previously described are applicable only to simple crystals, or to crystals whose physical and chemical properties are a great help in determining the structure. Most crystals are thus too difficult to yield to these methods, *e.g.*, the aromatic organic crystals, except a few isolated cases.

Another method of direct solution of crystal structure was developed by Ott.† With the assumption of spherically symmetrical atoms he obtains for a crystal having a centre of symmetry

$$\mathrm{F}\,(h00) = \sum_{1}^{n} \psi_r \,.\, e^{2\pi i \xi_a{}^r h/a}, \tag{2}$$

where ψ_r is the atomic structure factor and $\xi_a{}^r$ the x-co-ordinate of the rth atom.

If the atoms are identical we may put

$$\frac{\mathrm{F}\,(h00)}{\psi} = \mathrm{S}_h = \sum_r \beta_r{}^h, \tag{3}$$

where $\beta_r = e^{\frac{2\pi i \xi_a{}^r}{a}}$. If the n values of the β's be the roots of the equation

$$\Pi\,(x - \beta_r) = x^n + a_1 x^{n-1} + \ldots + a_n = 0, \tag{4}$$

then by Newton's relation, we have

$$\left.\begin{array}{l} \mathrm{S}_{100} + a_1 = 0 \\ \mathrm{S}_{200} + a_1\mathrm{S}_{100} + 2a_2 = 0 \\ \mathrm{S}_{300} + a_1\mathrm{S}_{200} + a_2\mathrm{S}_{100} + 3a_3 = 0 \\ .\quad\ .\quad\ .\quad\ .\quad\ .\quad\ . \\ \mathrm{S}_{n00} + a_1\mathrm{S}_{(n-1)\,00} + \cdots + na_n = 0 \end{array}\right\} . \tag{5}$$

Since the crystal has a centre of symmetry, we have (5),

$$\left.\begin{array}{l} 1 = a_n \\ a_1 = a_{n-1} \\ a_2 = a_{n-2} \\ a_3 = a_{n-3} \\ \text{etc., etc.} \end{array}\right\} . \tag{6}$$

---

* 'Proc. Roy. Soc.,' A, vol. 136, p. 210 (1932).

† 'Z. Kristallog.,' vol. 66, p. 136 (1928).

Thus we have a large number of equations, much more than required for finding the a's. Other similar equations may be formed for the other zones which we shall consider later.

Ott solved these equations for a few crystals, and the values of a's thus obtained were substituted in equation (4), whence the parameters were obtained directly. Only in exceptional cases is it possible to use this method, for the following reasons :—

(1) From X-ray diffraction it is possible to know the numerical values only of the S's and not their signs.

(2) These numerical values again are not generally sufficiently accurate to warrant a direct solution for the parameters.

(3) The equation (4) is one of the nth degree and hence in general cannot be solved if n is greater than 4.

Though these considerations preclude the possibility of a general use of this method, we can, however, find the signs of the Fourier terms with the help of these equations.

Let us consider the projection on the xy plane. Instead of trying to find the values of the a's let us eliminate them ; since there are a very large number of equations we can choose $\frac{1}{2}n + 1$ out of them, and by elimination we get a relation between the S's. We know the numerical values of these quantities and by trial we can find out the proper signs which will make them satisfy these equations.

But it will often happen that there will be more than one combination of signs that will satisfy these equations. This will happen more often if n is large, and then we have to find the a's from these combinations of signs and to substitute them in other available equations. By this means we can determine the combination that satisfies all the equations and also the signs of the other terms, for the (h00) planes. In a similar manner it is possible to find the signs of the terms corresponding to any plane, where measurements for a number of orders are available. The signs of the other terms may be obtained in the following way. We have to take the equation

$$\mathrm{S}\,(hkl) = \frac{\mathrm{F}\,(hkl)}{\psi} = \sum_r e^{2\pi i(\xi_a{}^r h/a + \xi_b{}^r k/b + \xi_c{}^r l/c)}$$

$$= \sum_r \mathrm{A}_r \gamma_r{}^h,$$

where $\mathrm{A}_r = e^{2\pi i(\xi_b{}^r k/b + \xi_c{}^r l/c)}$ and $\gamma_r = e^{2\pi i \xi_a{}^r h/a}$ and k and l are constants, h

being variable. For shortness we shall write S_h for $S(hkl)$. Let the n values of γ be the solutions of the equation

$$\prod_r (x - \beta_r) = x^n + b_1x^{n-1} + b_2x^{n-2} + \ldots b_n = 0.$$

On substituting these roots we have the following identities

$$A_1(\beta_1^n + b_1\beta_1^{n-1} + \ldots + b_n) = 0$$

$$A_2(\beta_2^n + b_1\beta_2^{n-1} + \ldots b_n) = 0.$$

$$\cdot \quad \cdot \qquad \cdot \quad \cdot$$

On adding we have

$$\Sigma A_r(\beta_r^n + b_1\beta_r^{n-1} + \ldots + b_n) = 0,$$

or

$$S_n + b_1 S_{n-1} + \ldots + b_n S_0 = 0.$$

Again the series of identities

$$A_r \beta_r^p (\beta_r^n + \beta_1 r^{n-1} + \ldots + b_n) = 0$$

give on addition

$$S_{n+p} + b_1 S_{n+p-1} + \ldots + b_n S_p = 0,$$

and similarly

$$S_{n-p} + b_1 S_{n-p-1} + \ldots + b_n S_{-p} = 0.$$

Thus we have a series of equations whose number depends on the number of intensity measurements available. With the help of the signs that have already been determined it is now possible to find the signs of these terms from these equations. It is not, however, possible to determine the signs of all the terms in this manner since the number of intensity measurements falls off with the order of the layer line. But when we have determined the signs of a number of terms we can proceed with the Fourier method and arrive at the correct structure after successive refinements.

Application to Anthracene.—As an example let us find the signs of the terms S_{00l} of anthracene crystal, for the different values of l. The number of carbon atoms in the unit cell is 28; but for every atom there is another having the same value of the z co-ordinate so we may take $n = 14$, the atomic structure factor for each atom being supposed to be doubled. As the crystal has a centre of symmetry, we have, according to equation (6),

$$c_8 = c_6; \; c_9 = c_5; \; c_{10} = c_4; \; \text{etc.},$$

where the c's correspond to the a's in equation (5), the number of unknown parameters being thus 7. So this is a fairly complicated case and nicely illustrates the method. The intensity measurements are taken from Robertson's* work. The signs found in this way agree well with those determined by Robertson from the rough structure obtained previously by the author.†

From the series of equations

$$\left.\begin{array}{l}
S_{001} + c_1 = 0 \\
S_{002} + c_1 S_{001} + 2c_1 = 0 \\
S_{003} + c_1 S_{002} + c_2 S_{001} + 3c_3 = 0 \\
\\
S_{007} + c_1 S_{006} + \cdots + 7c_7 = 0 \\
S_{008} + c_1 S_{007} + \cdots + 8c_8 = 0
\end{array}\right\} \qquad (7)$$

by elimination, we have

$$\begin{vmatrix}
S_{001} & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
S_{002} & S_{001} & 2 & 0 & 0 & 0 & 0 & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
S_{007} & S_{006} & S_{005} & S_{004} & S_{003} & S_{002} & S_{001} & 7 \\
S_{008} & S_{007} & S_{006} & S_{005} & S_{004} & S_{003} & (S_{002}+8) & S_{001}
\end{vmatrix} = 0. \qquad (8)$$

In Table I are given the values of " S " by dividing the structure factors from Robertson's results by the double of the atomic structure factors obtained from the curve given in Lonsdale's paper.‡

Table I.

Plane	S	Plane.	S.
000	14	006	0
001	2·8	007	0
002	2·5	008	0
003	2 6	009	3·0
004	4 2	0010	0
005	4·0		

Substituting the zero values and the numerical values where there are even powers of the S's in the equation (7), we have

$$\begin{aligned}
&- 4854 + 22849\, S_2 - 193\, S_1 S_3 + 5866\, S_4 + 2476\, S_1 S_5 \\
&- 2070\, S_1 S_2 S_3 + 14740\, S_2 S_4 - 4032\, S_1 S_2 S_5 - 3960\, S_1 S_3 S_4 \\
&+ 2688\, S_3 S_5 = 0.
\end{aligned}$$

---

* 'Proc. Roy. Soc.,' A, vol. 140, p. 79 (1933).

† 'Ind. J. Phys.,' vol. 4, p. 557 (1930).

‡ 'Proc. Roy. Soc.,' A, vol. 123, p. 494 (1929).

By trial it is found that the following three sets of signs, Table II, satisfy this equation.

Table II.

	1st set	2nd set	3rd set
S_{001}	+	+	+
S_{002}	−	−	−
S_{003}	+	−	−
S_{004}	−	+	−
S_{005}	−	+	−

Substituting these signs in the equations (6) we have the following sets of values for the c's in Table III.

Table III.

	1st set	2nd set.	3rd set
c_1	−2·8	2·8	2·8
c_2	5·2	5·2	5·2
c_3	8·1	−6·3	6·3
c_4	12·2	4·8	7·9
c_5	−15	1·6	−8·6
c_6	17	−1·8	6·4
c_7	−18·2	2·7	−3·5

On trying these sets of values in the equations

$$S_{009} + c_4 S_{005} + c_5 S_{004} + c_6 S_{003} + c_7 S_{002} + c_6 S_{001} + 9c_5 = 0$$

$$c_1 S_{009} + c_5 S_{005} + c_6 S_{004} + c_7 S_{003} + c_6 S_{002} + c_5 S_{001} + 10c_4 = 0,$$

it is found that only the first of these satisfy them, and the sign of S_{009} is found to be negative. The signs of the Fourier terms then agree with the signs found from the structure determined by the trial and error method. The process can be extended to determine the signs of the other terms as well.

We have taken the fairly complicated case of anthracene to exemplify the method, as, thanks to Robertson, a very good set of intensity measurements is available. Where n has a smaller value this method proves very convenient as the final equations are not so complicated. But the real usefulness of the method will be in rather complicated cases, where the method of trial and error fails, though a good series of intensity measurements has been made.

The author takes this opportunity to express his gratefulness to Dr. Robertson for allowing the author to use his unpublished data.

Collisions of α-particles with Fluorine Nuclei.

By N. FEATHER, Ph D., Fellow of Trinity College, Cambridge.

(Communicated by Lord Rutherford, O M , F R S —Received February 2, 1933)

[PLATES 1 AND 2]

Introduction.

Present knowledge concerning individual collisions of swift particles and nuclei is derived mainly from the researches of Blackett and his co-workers* , it has reference to the collisions of α-particles with hydrogen, helium, nitrogen, oxygen and argon nuclei—and of H-particles with oxygen nuclei. Its content is briefly as follows · conservation of kinetic energy and momentum has been found to hold, to the accuracy of measurement, in all save obviously disintegration collisions (disintegration has been found, amongst the above, only for nitrogen), and an empirical relation between velocity and range has been derived for each of the five types of recoil atom in question. The present paper describes an investigation of the collisions of α-particles with fluorine nuclei by the same general method ; it establishes the appropriate range-velocity relation, though it does not afford any positive information concerning the inelastic (disintegration) collisions which from other evidence are known to occur. Less complete data establishing the range-velocity relation for carbon recoil atoms have been obtained as a bye-product in the course of the work.

Fluorine has been employed in the present work very conveniently in the form of the saturated carbon compound, carbon tetrafluoride. In general the application of expansion chamber methods to fresh elements has been greatly retarded on account of the chemical activity of the simpler chemical compounds available. Of the lighter elements carbon has received less attention than might otherwise have been profitable presumably because it appeared unlikely that any disintegration effects would be observed, but, for the rest, boron, silicon, phosphorus, sulphur and chlorine, in order to be examined, must have been employed in the form of highly active gases. In such conditions technical difficulties arise in the construction of a suitable chamber and the finding of a convenient substitute to employ in place of the usual water vapour.* With elementary fluorine these difficulties are doubtless effectively insuperable; with carbon tetrafluoride, on the other hand, they entirely disappear. This compound has been reported at various times† in the past, but it remained for Lebeau and Damiens‡ in 1926-30 and Ruff and Keim§ in the latter year, definitely to establish its properties, both chemical and physical. The most important, in view of its employment in the expansion chamber may be summarized as follows: boiling point, $-130°$ C.; ratio of specific heats, C_p/C_v, 1·17, molecular stopping power, 2·75 (*v. inf.*); action on water, glass, copper, silver, lead, nil, very slow absorption by some oils Except in a fully equipped laboratory the preparation of the gas cannot be easily undertaken, since elementary fluorine is required. Professor Lebeau, however, kindly undertook to provide a sufficiently large sample for the work to be described. For this most generous gift of 2 litres of the gas I wish here to offer my sincere thanks

---

* Blackett, ' Proc. Roy. Soc.,' A, vol. 103, p. 62 (1923) ; vol 107, p 349 (1925) ; Blackett and Hudson, *ibid.*, vol. 117, p. 124 (1927), Blackett and Champion, *ibid.*, vol. 130, p. 380 (1931) ; Blackett and Lees, *ibid.*, vol. 134, p. 658 (1932), vol. 136, p. 325 (1932).

The Expansion Chamber and Camera System.

The expansion chamber was built in 1927 by the Cambridge Scientific Instrument Company to a design for which Drs Chadwick and Nimmo were chiefly responsible. The chamber described by Auger‖ in 1926 served as pattern. In 1929 the characteristics of the new apparatus were investigated by Nimmo and Richardson,¶ and the chamber itself was used by the latter worker in 1930 in an investigation of the β-particle emission of radium D, being very briefly described in the published account of the work.** Fig. 1, taken from the makers' drawings, shows the details of the design †† For the present investigation it became necessary to adapt the expansion chamber for automatic operation.‡‡ It would serve no useful purpose to enter into full details here, since the machine was not designed as a unit from start to finish; it incorporated large portions of the mechanical accessories of the apparatus of Nimmo and Feather* (1928) and of that of Nimmo and Richardson (*loc. cit.*), all, however, entirely rebuilt, with several additions.

---

* Progress towards the solution of the difficulty along these lines has, however, already been achieved, *cf.* Philipp, 'Z. Physik,' vol 53, p 100 (1929).

† Moissan, 'C R Acad. Sci. Paris,' vol. 110, p. 276 (1890), Humiston, 'J Phys. Chem.,' vol. 23, p. 572 (1919).

‡ 'C. R. Acad Sci. Paris,' vol. 182, p 1340 (1926), vol. 191, p 939 (1930)

§ 'Z. anorg. Chem.' vol. 192, p 249 (1930).

‖ 'Ann. Physique,' vol. 6, p 183 (1926)

¶ Nimmo, 'Ph.D. Dissertation,' Cambridge, 1929.

** 'Proc Roy Soc.,' A, vol. 133, p 367 (1931).

†† The actual drawing for this figure was made by Nimmo.

‡‡ The experiments here described were suspended in February, 1932, when work on the nuclear radiation from beryllium was begun Only incidental changes were made in the apparatus at that time. A short description of the arrangement was included in the initial report of the work, 'Proc. Roy Soc.,' A, vol. 136, p. 709 (1932).

A set of cams mounted on a common shaft and driven through reduction gears by a small electric motor controlled the cycle of operations. In normal working a complete cycle occupied from 28 to 32 seconds, according to circumstances—almost twice as long as employed by Blackett† for a chamber of very similar size. The controlling factor here is the mass of the piston (the stout construction of which is evident from the figure), a much greater time being required for full return before the main expansion than with a piston of

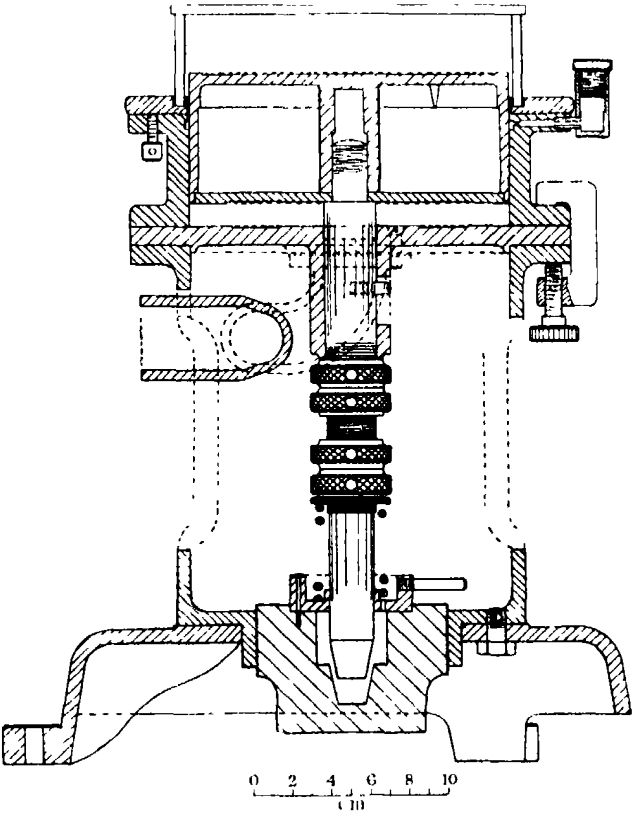


Fig. 1.—Cloud expansion chamber, Cavendish 1927 type.

---

* 'Proc. Roy. Soc.,' A, vol. 122, p. 668 (1929).

† 'J. Sci. Instr.,' vol. 6, p. 184 (1929).

smaller mass. A similar effect holds for the motion in expansion also; in order to keep this sufficiently rapid a specially large valve, of about 10 sq. cm. area, was employed. The opening of the valve was made as sudden as possible by using a light piston-like bung which was held in position until the instant of expansion and was then released, to be acted upon by the full pressure difference established between the space below the piston and the "vacuum reservoir." After a free travel of several centimetres the piston bung was brought to rest and, at an early stage in the subsequent cycle, was returned. For the intermediate clearing expansion extreme speed is not essential and an auxiliary valve of the usual pull-off type was employed.

The time for which α-particles were admitted to the chamber was controlled electromagnetically. At an instant which could be finely adjusted with respect to the time of expansion the current through a small electromagnet was broken In this way the slit system was released; a plate with a vertical slit in it moved horizontally across the source. Later, when the current was re-established, the slit system was automatically re-set. The source consisted of a rectangular platinum-faced button, 3·6 mm by 1·6 mm, activated by exposure to emanating radiothorium, and fixed with its long axis horizontal about 5 or 6 mm. from the mica window of the chamber The α-particle window was about 1 mm in diameter and was formed of mica of rather less than 1 cm. stopping power. A hole 7 mm in diameter had been made in the thick glass wall of the chamber. Into this was inserted a hollow plug of brass carrying the window in its outer face, and in its inner face a horizontal slit 1·5 mm. wide, intended further to limit the beam. The chamber had a diameter of 17 cm. and a maximum depth of 4·2 cm. By means of a second electromagnetic device similar to the first a light rod, carrying an insulated wire bridge at one end and pivoted near the other end, was in turn caught up, held, and released at the appropriate moment, in order to discharge a condenser (1/20 mfd. charged to 30 k.v) through two quartz mercury lamps in parallel and so provide the illuminating flash for the photographs.

The camera was constructed from the parts of two similar Ensign film cameras of an old design. These were built and coupled together so that two photographs were taken on the same length of film—each occupying the full width of the film—separated by such a distance (corresponding to a stereoscopic angle of 8° 54′) that six other full size photographs could be accommodated in the intervening space. In this way one lens was made to take always "even numbered" photographs and the other lens "odd numbered" photographs, when the film was moved forward two places between each exposure and the next.[*] The lenses were a matched pair of Cooke "Aviars" of 3⅝-inch focal length. In general they were used at maximum aperture, *f*/4·5. In the camera the film moved through guides in such a way that the image planes of the two corresponding photographs of a pair were suitably inclined[†] and the object planes, therefore, coincident—or nearly so[‡]—in the chamber itself. The film, in lengths of 100 feet each, was developed by Messrs. Kodak, Ltd., Kingsway, London. After development the photographs were numbered and examined. Finally, measurements were made on selected photographs by taking down the camera and using it in the stereo-reprojection apparatus of Nuttall and Williams,[§] as previously described.[‖] The ends of the tracks of the α-particles both of thorium C and of thorium C′ were visible on each photograph. In this way direct translation of measured lengths to reduced air ranges[¶] was always possible.

The Composition of the Gas Mixture.

On the basis of Bragg's rule the molecular stopping power of carbon tetrafluoride is 2·75 (air = 1). Moreover, under the conditions obtaining in the experiment to be described the α-particles of thorium C′ had a residual air equivalent range of somewhat more than 7 cm in the chamber. In order, therefore, that the whole width of the cloud volume should be used to the best advantage, it was necessary to dilute with a foreign gas—since the ideally preferable alternative of reduced pressure was altogether impracticable. In such circumstances helium is the most satisfactory dilutant—for two reasons. The molecular stopping power of helium is less than that of any other gas, and, since it is monatomic, smaller expansion ratios are possible, an advantage with a heavy piston. Now, if optimum proportions are to be calculated exactly, the ratio of the specific heats must be known for each component. For carbon tetrafluoride the appropriate data were not initially available. However, for methane we have $\gamma = 1{\cdot}31$, for methyl chloride 1·28, for methylene dichloride 1·22, for chloroform 1·15, and for carbon tetrachloride 1·13.* As a rough estimate $\gamma = 1\cdot17$ (14/12) was assumed for the corresponding tetrafluorine compound. On this assumption a mixture of 4 parts helium and 1 part carbon tetrafluoride (proportions by volume) is characterized by an effective value of γ of 1·42—almost exactly the value for a simple diatomic gas—and, at the end of an expansion for which $v_2/v_1 = 1\cdot36$, by an effective α-particle stopping power of 0·51.

---

[*] I wish here to thank Dr. F. R. Terroux for suggesting the basic idea involved in this arrangement.

[†] Blackett, 'Proc. Roy. Soc.,' A, vol. 123, p. 613 (1929).

[‡] Constructional limitations prevented the exact attainment of this condition, but the angle between the object planes was arranged to be sensibly smaller than that between the axes of the camera lenses.

[§] 'Proc. Phys. Soc.,' vol. 42, p 212 (1930).

[‖] Feather, 'Proc. Roy. Soc.,' A, vol. 136, p. 709 (1932).

[¶] Blackett and Lees, 'Proc. Roy. Soc.,' A, vol. 134, p. 658 (1932).

Such a mixture was made by filling the chamber with helium (crude helium containing 20 per cent. air was passed slowly over charcoal cooled in liquid air), taking out gas to a calculated pressure, and admitting the appropriate quantity of carbon tetrafluoride. The final amount of gas in the chamber was such as to allow of an expansion ratio of about 1·36 when the initial pressure was roughly atmospheric. When the test was made good tracks were obtained with this ratio, of almost the calculated length. It was concluded, therefore, that the values assumed, both for the stopping power and for the ratio of the specific heats, could not have been far from correct, and, as a result, the 4 : 1 helium-carbon tetrafluoride mixture was adopted for the whole course of the work. Even with these molecular proportions 66 per cent. of the stopping power of the mixture is provided by the atoms of fluorine; at this considerable dilution, therefore, two-thirds of the maximum possible yield of fluorine collisions should still be obtained.

In general, before the helium filling, the air in the chamber was replaced by hydrogen, so that, although 3 per cent. by volume of this initial filling remained, no other disintegrable element should have been present to an appreciable extent in the gas.

The Experimental Results.

Over a period of 5 months 6265 stereoscopic pairs of photographs were taken. These included the tracks of 455,000 α-particles—296,000 those of thorium C′,† under observation for the last 7·1 cm. of the air equivalent range, the remainder, the α-particles of thorium C to the extent of the last 3·2 cm. of path. No obviously disintegration collision was revealed by the mere examination of the photographs for purposes of registration and cataloguing.

Measurements were made on some 70 forked tracks, selected from a much greater number actually recorded. Simply by inspection it was usually possible to decide whether an α-particle-helium collision was in question, or whether the collision had been with either a carbon or a fluorine nucleus. Very few measurements were made on forks of the former (helium) group, and those merely as a guide to the general accuracy of the determination. In the latter group all collision photographs in which the reduced air range of the recoiling particle was greater than about 2 mm. were submitted to examination in the reprojection apparatus, and only a very few of these discarded as unprofitable for measurement. Amongst those forks with shorter recoil spurs the choice was necessarily somewhat arbitrary, since in all 987 forks of the "C or F" type were catalogued (180 helium collisions were observed).

---

* Capstick, 'Proc. Roy. Soc.,' A, vol. 57, p. 322 (1895); 'Phil. Trans.,' A, vol. 185, p. 1 (1894).

† This estimate is based on representative counts in which 4528 separate tracks were counted.

Measurements on the helium forks consisted in single determinations in five instances of the angle between the two arms of the fork. In two $(\theta + \phi)$ was directly determined, and in the remaining three cases θ and ϕ were measured individually. Reproductions of four of the forks in question are included in Plate 2. The following values were obtained :—

Plate 2, No	14	16	17	20	*
$\theta + \phi$	91·5°	89·8°	93·0°	87·9°	85·5°

* Not reproduced

On account of the equality of mass of the two particles involved in the collision the value to be expected in each case is 90° Allowing a small margin of error this is precisely what is found, moreover the mean of the determinations above quoted is 89·5°, in good agreement with the theoretical value.

A single determination was also made of the angles of deflection (ϕ) and projection (θ) for one collision, very obviously between an α-particle and a hydrogen nucleus. A reproduction of one photograph of this pair is given in Plate 2, No. 22. The measured angles were as follows: θ 41·2°, ϕ 15·2°, corresponding to a mass ratio of 0·265 instead of 0·252. It is almost certain that the true value of ϕ was 14·4°† —the maximum deflection possible (when the true value of θ is about 40°). By making the necessary calculations for this collision a single point on the range-velocity curve for fast protons was obtained. The point in question is r (reduced air range) $= 17 \cdot 9$ mm, $v = 12 \cdot 1 . 10^8$ cm./sec. From the results of Blackett and Lees (*loc. cit.*) and Blackett,‡ corresponding to this range we should expect a velocity of $12 \cdot 4 \, . \, 10^8$ cm./sec. The agreement is again satisfactory.

---

† Blackett and Lees, 'Proc. Roy. Soc.,' A, vol. 134, p. 658 (1932).

‡ 'Proc. Roy. Soc.,' A, vol. 135, p. 132 (1932).

In general only a single set of measurements was made of lengths and angles for each of the selected forks classified as "C or F" (*v sup.*). Sometimes, however, two or more independent settings of the film were made and the corresponding measurements recorded. The values given in Table I show the degree of consistency attained. Here l_1 and l_2 are the reduced air ranges, after the collision, of α-particle and recoil atom, respectively, and θ and ϕ have their usual significance

Table I.

Serial No	θ	ϕ	l_1 cm	l_2 mm	Reproduction (if any)
II 646	26·4 24·3	98·9 97·4	—	3·32 3·24	—
III 583	51·2 49·8	51·9 52·5	2·57 2·67	2·07 2·18	-
III 656	17·2 16·2	132·4 131·5	- —	7·46 7·35	Plate 1, No 7
V 433	54·7 57·4	68·0 67·8	$>1 \cdot 06$ $>1 \cdot 08$	3·14 3·20	—
VI 453	42·5 44·6	86·1 82·7	—	3·62 3·56	Plate 1, Nos 12, 13

In each of the examples in Table I a fairly long recoil track was available for measurement, and the accuracy of determination of θ was not very inferior to that of determining ϕ. It is otherwise when the shorter recoil tracks are in question. On the assumption that in fixing the initial direction the full length of each recoil track was used—to do which was frequently quite impracticable—and that the probable resultant transverse positional error involved in setting on the two ends was 0·2 mm. in the full size image, it was estimated that the average value of the probable error in θ, taken over all the collisions investigated, was 3·6°. The individual values from which this average was drawn included 7 cases of a probable error less than 1·5° and 4 cases with an error greater than 7°. Now in the majority of practical cases the calculation of the mass ratio for the colliding particles is limited by this somewhat inaccurate knowledge of θ, the angle of projection of the struck nucleus, especially if the latter be heavy. It has been pointed out above that this is the more difficult measurement; it will now be shown that it is often the more important. For, if m be the mass of the nucleus and M the mass of the α-particle

$$\frac{m}{\mathrm{M}} \; (\equiv \gamma, \text{say}) = \frac{\sin \phi}{\sin (2\theta + \phi)} \tag{i}$$

Thus

$$\sin (2\theta + \phi) \,.\, \delta\gamma = 2\gamma \left[\frac{\sin 2\theta}{2 \sin \phi} \,.\, \delta\phi - \cos (2\theta + \phi) \,.\, \delta\theta \right] \tag{ii}$$

Now it may easily be established that for $\gamma \nless 3$, namely, for collisions with all nuclei heavier than that of boron, B^{11}, for all possible values of θ and ϕ,

$$| \cos (2\theta + \phi) | > | \cos \theta |,$$

Moreover, in more than 60 per cent. of the C or F collisions investigated, $\sin \phi > \sin \theta$. Thus, for this majority at least, and presumably for some others also, must

$$| \cos (2\theta + \phi) | > \left| \frac{\sin 2\theta}{2 \sin \phi} \right|,$$

and so the corresponding measurements of θ carry more weight in the determination of γ than those of ϕ [equation (ii)]. As a numerical example of the magnitude of the uncertainty, if $\gamma = 4\cdot75$ (fluorine collisions), $\delta\theta = 3^\circ$, $2\theta + \phi = 168\cdot7^\circ$ ($\cot^{-1} 5\cdot0$), then, neglecting any error arising from a possible inaccuracy in ϕ, $\delta\gamma = 2\cdot5$. To distinguish between α-particle-carbon and α-particle-fluorine nucleus collisions turns out, therefore, to be generally impossible, at least on the sole basis of individual measurements. If ten independent settings of the film had been made, together with appropriate measurements, for each pair of collision photographs, more trustworthy decisions might probably have been reached. It was not thought worth while to attempt this standard of precision. For, to a first approximation, and in the domain of normal Coulomb scattering, fluorine collisions are nine* times more probable than the corresponding carbon collisions, for the gas mixture employed. Furthermore, the range-velocity relations are such that the assumption of a mistaken value for γ in any case has roughly the effect of moving an experimental point from its true position to a position very close to the true curve appropriate to the mistaken value of γ. For an accurate determination of the range-velocity relation for carbon recoil atoms a more suitable gas mixture must eventually be employed; by contrast the presence of carbon in the present mixture cannot give rise to appreciable error in the fluorine results.

---

* *I.e.*, $4 \times (9/6)^2$.

The values of r_n and v_n finally calculated for the recoil atoms from 63 collisions are shown by the points in fig. 2. In 55 cases $\gamma = 4 \cdot 75$ was adopted, the more probable assumption concerning the nature of the collision being accepted as correct. In the remaining cases $\gamma = 3$ was preferred on the grounds that the experimental value deduced from (i) was less than 4·75 by more than the amount calculated for the partial probable error owing to inaccuracy in θ. This is admittedly rather an arbitrary procedure. Moreover, it leads to a surprisingly large fraction of the longest recoil spurs being attributed to the recoil atoms of carbon—though in this, perhaps, it receives some measure of justification, for it is here that anomalous scattering is becoming important, and more markedly so the smaller the atomic number of the scattering nucleus. For each point it may be seen by inspection of the figure whether the velocity of recoil has been obtained from the assumed value of γ and the measured length of path of the α-particle after collision (through the corresponding final velocity, v_α, by equation (iii)) or, if the path of the deflected

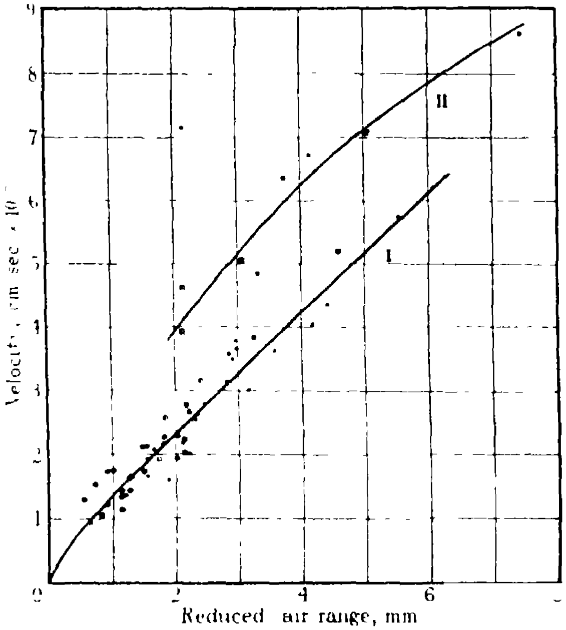


Fig. 2.—Curve I, fluorine recoil atoms; curve II, carbon recoil atoms. ⊙ from v_α, with $\gamma = 4 \cdot 75$; ● from V, with $\gamma = 4 \cdot 75$; ⊡ from v_α, with $\gamma = 3 \cdot 00$; ■ from V, with $\gamma = 3 \cdot 00$.

particle has not been completely observed, from γ and the estimated residual range of the α-particle at the point of collision (through V, the original velocity, by equation (iv)). In terms of γ and the experimentally determined quantities the equations in question are

$$v_n \gamma \sin \theta = v_a \sin \phi \tag{iii}$$

$$v_n (1 + \gamma) = 2V \cos \theta. \tag{iv}$$

It would appear from the results that these two methods of deducing the velocity are affected by roughly comparable errors, though the process of selecting observational material has resulted in the former being chiefly used for the shorter, the latter for the longer recoil tracks. Both have been given the same weight in deducing the final range-velocity curve (I). Curve II is an indication, from the meagre data available, of the form of the range-velocity relation for carbon recoil atoms of moderately high speed.

Discussion.

From preliminary results,* in 1925,† Blackett suggested the general relation

$$r = kmz^{\frac{1}{2}} f(v), \tag{v}$$

by which the range r of a recoil atom of mass m and atomic number z is simply related to its initial velocity v The constant k and the form of the function f were assumed the same for all atoms. It is now known—as the following examples testify—that the above expression is of very restricted applicability. The ranges of fast protons and α-particles are roughly the same for equal velocities ($> 10^9$ cm./sec.), for a velocity of 5.10^8 cm./sec. the ranges of α-particles and nitrogen recoil atoms are very nearly equal, and at $8 . 10^7$ cm./sec. α-particles and argon recoil atoms have the same range. Yet by equation (v) between corresponding ranges in these cases a factor of the order of 2 or more is to be expected. On the other hand there are some surprising successes in applying this relation. As a notable example it may be remarked that in the region 3–4 . 10^7 cm./sec. the ranges of the radioactive recoil atoms‡ are from 2 to 2·5 times as great as those of atoms of argon recoiling with the same velocities, whereas the ratio between ranges to be expected from (v) is 2·4. Furthermore, though the range-velocity curves for nitrogen and argon recoil atoms are different in form, yet in each there is a well-marked linear portion, and if the slopes of these sections* only be considered an experimental ratio of 1·82 is obtained, where the above formula predicts 1·78. So close an agreement is doubtless fortuitous, but at least it establishes the usefulness of the formula in selected cases.

---

* 'Proc. Roy. Soc.,' A, vol. 103, p. 62 (1923).

† 'Proc. Roy. Soc.,' A, vol. 107, p. 349 (1925).

‡ Wood, 'Phil. Mag.,' vol. 26, p. 586 (1913); Wertenstein, 'Ann. Physique,' vol. 1, pp. 347, 393 (1914); Kolhörster, 'Z. Physik,' vol. 2, p. 257 (1920).

It will be seen that curve I, fig. 2, here established for the recoil atoms of fluorine, is almost of the same form as the curves for oxygen and nitrogen recoil atoms given by Blackett and Lees. Equation (v) predicts a ratio of ranges, $r_{F^{19}}/r_{N^{14}}$, at a fixed velocity, of 1·20; the ratio of the slopes of the experimental curves is 1·29 The discrepancy is a little greater than has been noted in some ratios ($r_{O^{16}}/r_{N^{14}}$, $r_{O^{17}}/r_{N^{14}}$), but it is not outside the combined limits of error of the two sets of measurements As between $r_{C^{12}}$ and $r_{N^{14}}$ equation (v) predicts 0·93; taking the results of Blackett and Lees for $r_{N^{14}}$, this gives for $r_{C^{12}}$ almost exactly the slope of curve II at the lower limit of its determination Equally well, however, the last point on this curve establishes beyond doubt the fact that the range of the carbon recoil atom varies with a higher power of the velocity as the velocity is increased.†

At this stage it is convenient to examine somewhat in greater detail the "range" employed in constructing the curves under discussion. In the conditions obtaining in the various experiments the recoil atom loses energy in encounters with the atoms of a gas mixture the stopping power of which may be chiefly (as with fluorine), or to a negligible extent (as for the radioactive recoil atoms, or for O^{17} produced in disintegration collisions), caused by atoms of the same type as the atom in motion. The actual range is measured, and the reduced air range calculated on the assumption that the stopping power of the gas mixture for the recoil atom is the same as the mean stopping power for complete absorption of α-particles of a certain specified initial velocity. There is no *a priori* reason why this should be so. Loss of velocity of recoil atoms must to a large extent be the result of the capture and loss of electrons, a process in which an atom of the gas and the recoiling atom are equally involved. To obtain quantitative results theoretically it will obviously be necessary to use our knowledge of the electron states of the two atoms. Physically, however, this much may be said, that, where for each pair of colliding atoms it would seem that a separate calculation were required to determine the rate of exchange of energy, it is very surprising that some approach to generality may be reached with a simple formula in which the recoil atom is characterized by mass and charge numbers only, and the properties of the struck atom do not enter explicitly. It may be that more exact research—in which ranges are determined in a standard gas—will establish complications which have not hitherto appeared. Possibly a regularity with the period of the Mendeléeff table will be found as the most general expression of dependence on the outer electron structure of the atom. It is in this connection that extreme cases, such as those of fluorine and argon, are of interest Moreover, if intense sources of homogeneous neutrons become available in the future, it may be possible to determine range-velocity curves over whole periods of consecutive elements in the table.

---

* $v > 1·5 . 10^8$ cm./sec. for argon, $v > 2·5 . 10^8$ cm./sec. for nitrogen recoil atoms.

† If the fork here in question be ascribed to any other possible collision (*e.g*, with nitrogen, oxygen or fluorine) then a similar change of slope in the corresponding range-velocity curve is with certainty established.

The present results may be discussed from a further point of view. Chadwick and Constable,* have shown that the yield of disintegration collisions in pure fluorine is about 3·8 per million α-particles from polonium. Assigning an effective fluorine content of 66 per cent to the gas mixture here employed, a probable yield of 1·07 such inelastic collisions was to be expected in the course of the work. That no single example was found gives little cause, therefore, for undue surprise. The above calculation, however, neglects the greater efficiency of faster α-particles, and it might well be that nearly 300,000 α-particles, under observation from 7·1 cm residual range to 3·81 cm. the mean range of the α-particles from polonium, should contribute at least another disintegration collision. Yet no such collision was observed. The disintegrations produced by polonium α-particles for the most part appear to result from capture through resonance levels of positive energy, and considerations of sharpness of resonance indicate, in the two best established instances, that the angular momentum of the effective α-particle about the capturing nucleus is between $h/2\pi$ and $2h/2\pi$. This last result depends on what efficiency is assumed; whether or not the probability of capture approaches unity at the centre of the level when the angular momentum condition is fulfilled. However, except statistically, it is difficult to obtain evidence on this point. For individual elastic collisions track measurements fix the residual range, and so the velocity of the α-particle at the instant of collision—within the limits which straggling inevitably imposes—and from the measured angles calculation of the impact parameter of classical mechanics is possible on the assumption of a Coulombian scattering field. But this impact parameter, and the corresponding angular momentum, have no unique significance in wave mechanics, moreover, departures from an effectively Coulombian field are known to occur, whilst at certain resonance energies more complicated conditions obtain. In actual fact no collision at either resonance energy was observed for which the classically calculated Coulomb field impact parameter corresponded to 1–2 units of angular momentum as above specified. On one occasion, however, No. 5, Plate 1, a pseudo-central collision occurred elastically at an α-particle energy within the limits associated with the longest range groups in the proton spectrum Now, if these disintegrations also result from resonance capture, central collisions are presumed to be responsible * Two other pseudo-central collisions were recorded, in these the initial energy was almost certainly greater than the energy corresponding to the peak of the potential barrier Here the domain of anomalous scattering has certainly been reached and no definite interpretation is possible.

---

* 'Proc. Roy. Soc.,' A, vol. 135, p. 48 (1932).

But disintegration with the emission of a proton is not the only type of disintegration to which the fluorine nucleus is susceptible Chadwick† has shown that neutrons may also be obtained, though in smaller numbers, when the α-particles from polonium are employed. The inelasticity of this type of collision would chiefly appear as a lack of coplanarity in the visible components of the fork. On account of the relatively small momentum of the particle which is liberated this discrepancy would not be expected to be considerable; to establish the effect with certainty must therefore require a large amount of experimental material. All that can be done on the basis of the data here obtained is to record the fact that in one case an apparent lack of coplanarity amounting to about 26°‡ was observed. Unfortunately the complete path of the α-particle after collision was not obtained, so that no satisfactory calculation of momenta can be made. The energy of the α-particle before collision was about $5 \cdot 5 \,.\, 10^6$ electron volts, not very different from the peak value in the potential barrier of the fluorine nucleus. It is to be expected on general grounds that the type of disintegration is not dependent upon the mode of penetration of the nuclear structure, and it will be interesting to see whether more detailed research exhibits the same resonance levels as effective whether proton or neutron is eventually expelled

---

* Chadwick and Constable, *loc. cit.* The elastic collision in nitrogen reproduced by Blackett and Lees, 'Proc. Roy. Soc.,' A, vol 134, p. 658 (1932), Plate 11, 3 is pseudo-central. The initial energy in this case, also, coincides approximately with a resonance energy, $3 \cdot 8 \,.\, 10^6$ electron volts—*cf.* Blackett and Lees, 'Proc. Roy. Soc,' A, vol. 136, p. 325 (1932); Pollard, 'Ph.D. Dissertation,' Cambridge (1932).

† 'Proc. Roy. Soc.,' A, vol 136, p. 692 (1932); later results in the press

‡ The angle between the stem and the plane of the forward visible branches of the fork

In addition to acknowledgments already made in the course of the paper, I wish here to express my appreciation to Professor Lord Rutherford and Dr. J. Chadwick for much helpful counsel throughout the experiment, and my gratitude to the Council of Trinity College for a grant from the Rouse Ball Research Fund which made possible the construction of the stereo-camera and provided other material for the work.

Summary.

Stereoscopic photographs have been taken of the tracks of nearly half-a-million α-particles in a mixture of carbon tetrafluoride and helium. The source (thorium B + C + C′) was placed outside the expansion chamber, with the result that the initial 1·5 cm. air equivalent range of each group of particles was not observable. Over the remaining portion of path no proton disintegration was obtained. The range-velocity curve for fluorine recoil atoms has been constructed from stereo-reprojection measurements on 63 selected collision photographs, check measurements being similarly made for five close collisions with helium nuclei and one with the nucleus of a hydrogen atom. From the main data eight cases were excluded as probably referring to collisions with carbon nuclei; from these some information concerning the form of the corresponding range-velocity curve has been obtained. A comparison is finally made with previous results and reproductions of a number of collision photographs are appended.

DESCRIPTION OF PLATES.

PLATE 1 —Collisions with carbon and fluorine nuclei. Magnification from standard r (approximate): No. 1, 0·75, Nos. 2–13, 2·5.

No 1.—Two well-marked spurs and the track of one long range α-particle on the same photograph

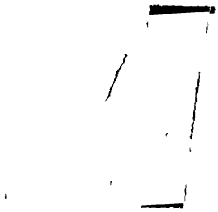
Nos. 3, 4, 7.—Presumably carbon recoil tracks, that of No. 7 7·46 mm. in length (reduced air range).

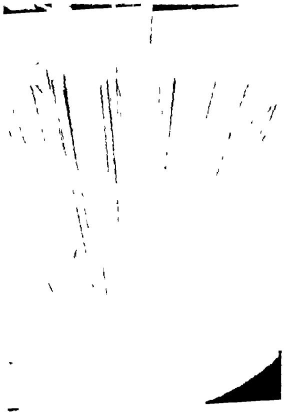
Nos. 2, 5, 6.—Large angle collisions with fluorine nuclei.

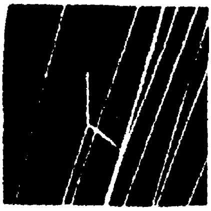
Nos. 12, 13.—A stereoscopic pair: fluorine collision.

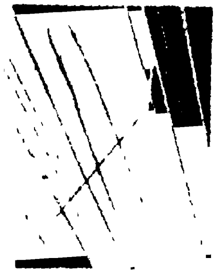
Nos. 10, 11 —A stereoscopic pair: peculiar collisions towards the end of the range.

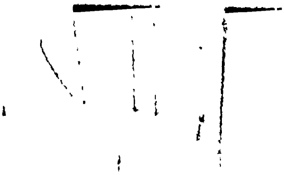
Nos. 8, 9 —Deflections with almost negligible spurs, in unusual associations.


2

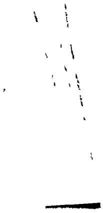

1


3


5


4

6


8

7

9

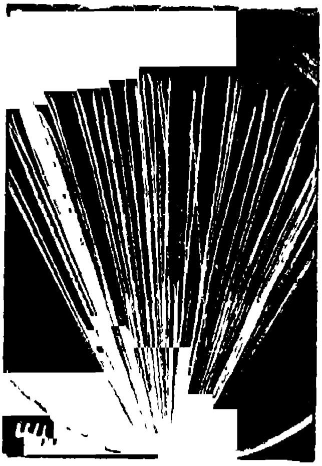
12

10

11

13

(*Facing p.* 208.)

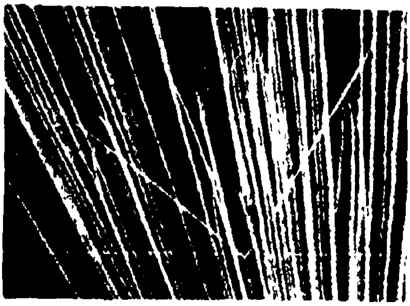


14

15

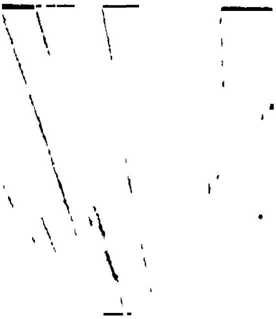
16

18

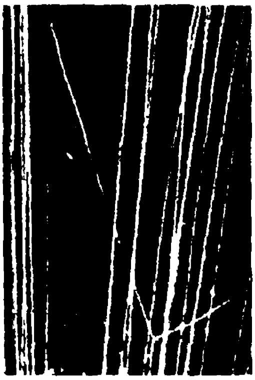


17

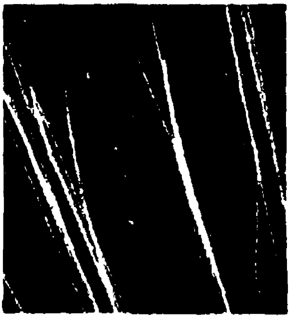
19



21

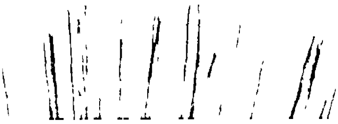


20

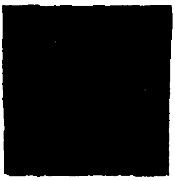


22

24



23



25

Plate 2.—Collisions with helium and hydrogen nuclei. Magnification from standard air. No. 14, 0·75; No. 23, 1·45, Nos 15–22, 24, 25, 2·5

No 14.—Complete photograph, showing both groups of α-particles, with helium collision almost in the plane of the chamber

Nos. 15, 16, 17, 21.—Helium collisions with $\phi \doteqdot \theta$, fast α-particles.

Nos. 18, 19, 24.—Helium collisions with $\phi \doteqdot \theta$, slow α particles.

Nos. 20, 25.—Unsymmetrical helium collisions.

No 22.—Collision with hydrogen nucleus $\phi \doteqdot 14{\cdot}4^\circ$.

No. 23.—Three nearly symmetrical helium collisions on one photograph—*cf.* Blackett and Champion, 'Proc. Roy. Soc.,' A, vol. 130, p. 380 (1931).

Internal Photoelectric Absorption in Halide Crystals.

By R. W. Gurney, Ph.D., Lecturer in Manchester University.

(Communicated by R. H. Fowler, F.R.S.—Received February 10, 1933.)

The possession of absorption bands in the ultra-violet by otherwise transparent crystals has been recognized for a long time as determining the dispersion of the medium. This absorption of radiation is one which produces no progressive change in the crystal On the other hand, the illumination of crystals which gives rise to photochemical and photo-conduction phenomena causes a progressive change in the crystal and consequently in its absorption spectrum. Measurements have recently been made of the original and induced absorption spectra of many photo-conducting crystals, in the hope of elucidating the formation of the latent photographic image

§ 1. Quantum mechanics gives us entirely new ideas as to how to correlate the absorption spectra of crystals with their insulating properties. Both phenomena are alike concerned with the quantized electron levels of the crystal. These allowed levels belong not to individual atoms or ions but to the crystal as a whole. The wave functions representing the valence electrons are not localized at individual lattice points, but are oscillatory throughout the crystal. Even in an insulator they are the wave functions of an electron gas moving in the periodic field of the lattice. The opacity of a metal for all wave-lengths from the infra-red far into the ultra-violet is due to the fact that electron transitions are possible to vacant levels of all higher energies. For other pure crystals, however, the allowed bands of levels are broken up by wide zones of disallowed energies, and it is these zones which at the same time give to a crystal the properties both of transparency to light and non-conduction to electric current. More exactly, these properties are due to the fact that a zone of forbidden energies divides a band of levels, which at low temperatures is completely filled with electrons, from a similar higher band which is completely empty. For crystals which are transparent to all visible light we can at once say that the width of this particular zone must be more than 3 ε-volts; for the alkali halides which are transparent into the Schumann region it must be at least more than 5 or 6 ε-volts. When illuminated with ultra-violet light of the right wave-lengths, however, such a crystal is almost on a par with a metal, it contains as many electrons, and the levels are similar; over narrow regions in the ultra-violet, therefore, it absorbs as if it were a conductor and shows selective metallic reflection.* The theory of the insulating properties of these crystals has been discussed by A. H. Wilson.† The same scheme of levels must be used for correlating conductivity with opacity, and insulation with transparency in the case of liquids, such as mercury and water.

A recent addition to the theory of the perfect lattice has been made by Tamm.‡ Investigating the wave functions for electrons, he has shown that at the boundary of a perfect lattice, there will be localized electron levels whose energies differ from those of the allowed bands of levels of the interior. That is to say, these levels lie in the disallowed zones of energy. When we place a crystal between metallic electrodes, electrons from the metal will fall into these surface levels charging the surface of the crystal negatively. This, Tamm suggests, must be why we are unable to send a stream of electrons through the empty bands of levels of the crystal, which otherwise we should be able to do.

Several pieces of recent experimental work have added to our knowledge of the permanent changes brought about in crystals by the absorption of light. The interpretation of these experiments is, however, being made in an older terminology very different from that outlined in the preceding paragraphs. Nor has there been sufficient recognition of the distinction between true absorption and the scattering which also removes light from the beam. In the following sections the various inter-related phenomena will be discussed in terms of transitions between the allowed electron levels. For this purpose the model of a perfect lattice must be modified to represent an imperfect crystal. We can at once say that at every submicroscopic crack in the interior there will be, on and near the surfaces, localized levels similar to those which Tamm placed on the external boundary, with energies lying in the disallowed zone. We shall have everything from definite submicroscopic cracks, grading down to small irregularities in the lattice which will give levels straying only slightly from the regular band. In practice it is found that even in good crystals the details of the ultra-violet absorption bands vary from one specimen to another.* Before attempting to base an interpretation on these levels, it will be necessary to review a few of the phenomena and their current interpretations, to see how they will fit into the proposed scheme of levels.

---

* Hilsch and Pohl, 'Z. Physik,' vol. 57, p. 149 (1929).

† 'Proc. Roy. Soc.,' A, vol. 133, p. 458; vol. 134, p. 277 (1931).

‡ 'Phys Z. Soviet Union,' vol. 1, p. 733 (1932).

§ 2. In investigating the permanent changes in crystals, it has always been the aim to avoid surface effects and to obtain a true volume effect. For this purpose the characteristic frequencies of the crystal which show metallic reflection, are useless. It has always been necessary to work with wave-lengths with an absorption coefficient, not of 10^5 cm.$^{-1}$ but of 1 cm.$^{-1}$ or less In the absence of impurities absorption coefficients of this magnitude are provided by the irregular levels. Smekal and his pupils have shown how the intensity of the latent image varies with mechanical strain, etc.† It is unnecessary, then, to discuss here the question whether light travelling through a perfect lattice at low temperature would produce any permanent change except at the boundary.

Owing to the difficulty of working on the short wave side of the band of intense absorption, radiation at the long wave side only has been used. Thus the quanta have abnormally small $h\nu$ and the transitions are between levels which are closer together than those responsible for the band of intense absorption. With these wave-lengths we have no metallic reflection. Instead we have to distinguish between (*a*) scattering of light, and (*b*) true absorption in which an electron suffers a permanent change of level. The latter causes a new absorption spectrum in the visible region to develop (colouration) Converting frequencies into ε-volts, the extent of the strongest part of this acquired absorption is as follows.

Table I.

	ε-volts			ε-volts	
NaBr	< 1·8 to	> 3·0	KBr	< 1·7 to	> 2·6
NaCl	< 1·9	> 3·5	KCl	< 1·9	> 2·7
NaF	< 2·6	> 4·2	AgCl	< 1·5	> 3·1

---

* Hilsch and Pohl, 'Z. Physik,' vol. 68, p. 724 (1931).

† "Atti. Cong. Int. Fis Como" (1927), "Int. Kong Photographie Dresden" (1931).

The interpretation usually given to this colouration process is that by absorption of each quantum an electron has been handed back from a halogen negative ion to a positive metal ion to give a neutral metal atom; and that it is the presence of these metal atoms in the crystal which is responsible for the acquired absorption. But in these cubic lattices the small metal ion is wedged in between six halogens, and occupies a volume about *one-third* of the volume of the much larger neutral metal atom; so that, if an electron is introduced into this small region, it will disturb all the surrounding electrons, and any stable levels can only come about as a property of the whole configuration. To ascribe the acquired absorption spectrum to the presence of neutral metal atoms seems at least an unsuitable terminology, especially as we have a number of different absorption spectra for a given metal, as seen for sodium in Table I.

It was formerly thought that in the alkali halides the only light capable of producing this colouration were the wave-lengths near the ultra-violet absorption bands; and this gave rise to the idea that the formation of latent image was a photochemical process requiring an almost definite quantum of energy.* But it has since been found† that in KBr at room temperature an acquired absorption in the red can be produced by illumination with any wave-length from 2030 A. to beyond 4360 A (that is, from 6 ε-volts to less than 2·8 ε-volts). This seems to quash the idea of a definite quantum process as essential to the change brought about.

Perhaps the most remarkable feature of the colouration of alkali halides is that it can be produced by warming a transparent crystal either in the metal vapour or in the molten metal. Atoms of the metal apparently diffuse rapidly into the interior through the submicroscopic cracks and glide-planes. On spectrophotometric examination it is found that the crystal has developed an absorption in the visible indistinguishable from that produced by light or X-rays.‡ This speaks in favour of the new absorption being due to electrons at the surfaces of cracks, since we cannot suppose that each foreign atom succeeds in mimicking the configuration of an atom in the interior of the lattice.

§ 3. These few considerations will have sufficed to show how difficult has been the interpretation as long as we had no sound basis on which to build. The best that one could do was to ascribe the absorption bands to the various ions and atoms in the crystal. For example, the absorption of KI at 2050 A. was assigned to the iodine negative ion. Thus, when a mixed crystal of KCl containing 1 per cent. of KI was made, it was anticipated that this absorption band of the iodine ion would be found.* The crystal, however, turned out to be transparent in this region. Now, on the other hand, by the use of quantum mechanics we have a definite picture of what determines the allowed levels. Since the wave functions extend throughout the crystal, it is the periodic potential of the lattice which determines the allowed energies of both the characteristic and irregular levels.

---

* Hilsch and Pohl, 'Z Physik,' vol. 77, p 423 (1932).

† Hilsch and Pohl, 'Z. Physik,' vol. 68, p. 728 (1931).

‡ Gyulai, 'Z. Physik,' vol. 37, p. 880 (1926).

It is well known that atoms of small ionization potential (alkali atoms) incident on a metal surface give up their electrons to the metal. Further, a metal is supposed to be able to give up electrons to the surface levels on a crystal; which implies that the energies are such that alkali atoms could give up electrons to these surface levels in the same way. This has interesting consequences if one accepts the view that the colouration of crystals is due to the presence of electrons in similar levels on the surfaces of internal submicroscopic cracks. Electrons may have been raised into these levels from lower levels, or may have been introduced from outside. For the latter purpose exposure to electrons escaping from a thermionic filament would, of course, be useless, since appreciable diffusion of electrons into the crystal would be prevented by their mutual repulsion. Positive charges must be introduced at the same time in order to obtain a considerable concentration of electrons. On this view, in the colouration of crystals by diffusion of metal, the metal atoms would be merely acting as the carriers of electrons. If this were so, any atom of sufficiently low ionization potential would do equally well, and we should have the unexpected result that the absorption spectrum would be independent of the metal used. Now it was found by Gyulai (*loc. cit.*) in 1926 that KBr coloured by means of sodium developed an absorption spectrum in the visible indistinguishable from that of KBr coloured by potassium; both spectra were identical with that produced by the action of ultra-violet light or X-rays. To account for this result Gyulai put forward the tentative explanation that the diffusing sodium atoms each replaced a potassium ion in the lattice, thereby setting free a potassium atom to act as an absorbing centre for visible light. This explanation betrays the universal tendency to explain absorption in terms of atoms rather than of electron levels.

The giving up of their electrons to the crystal by the diffusing metal atoms is further suggested by an experiment of Stasiw,[*] who has worked with crystals coloured by means of metal vapour. When such a crystal was placed in an electric field, and then warmed, the colouration was seen to move towards the positive electrode. On reversing the field the colouration could be moved in the other direction at will.

---

* Hilsch and Pohl, 'Z. Physik,' vol. 57, p 152 (1929); Smakula, 'Z Physik,' vol 45, p 1 (1927)

§ 4. Since the preceding sections were written an article by Wilson has appeared,[†] outlining a theory of photoconduction to embrace idiochromatic and allochromatic crystals. Both are there treated as occurring in an initially perfect lattice. The details of the model proposed here will consequently differ in many respects. The potential energy curve for the removal of an electron from any point O in a regular lattice will be roughly of the form

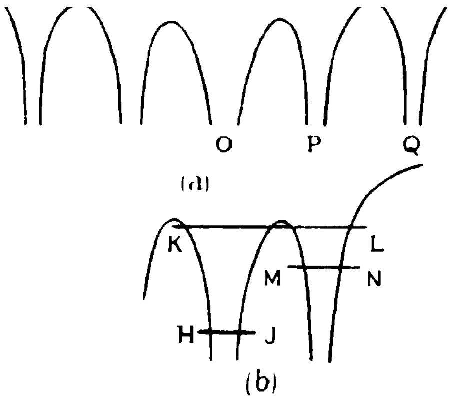


Fig 1

of fig. 1, *a*, the periodic potential being superposed on the mutual energy of the electron and its positive "hole," whose presence causes the local modification of the regular periodic potential shown in the diagram. If next we suppose, as in fig. 1, *b*, that the ion Q is absent, the ion P lying on the surface of a submicroscopic crack, the electron configurations between P and O and their neighbours will be different from those of normal ions close-packed within the lattice. Let MN be a localized level due to the presence of the crack, and initially empty. Let KL be another unoccupied level, and HJ an occupied level; we may keep an open mind as to whether either of these lies within a band of the lattice levels or not. We have to consider the action of light on an electron at HJ (or in a similar level in an ion not adjacent to P); the electron may either scatter a quantum and return to HJ, or else it may be transferred to MN, which is the colouration process.

---

[*] 'Nachr. Ges Wiss. Göttingen,' p. 261 (1932).

[†] 'Nature,' vol. 130, p 913 (1932).

Although a quantum is absorbed in the formation of this latent image, another quantum is absorbed for the decolouration process. This suggests that in the latter the electron is being lifted over a potential barrier before it can fall to a low level. It will appear that this is also true of the original transition to MN which is under discussion; that is, the latent image is not formed by a direct transition. It has been pointed out to me by Professor Fowler that since the fixed ratio of the Einstein A and B coefficients for any simple system is determined by a purely thermodynamic argument, there is no possibility of having B large and A very small, which would be necessary to enable an electron to stay almost indefinitely in the level MN. We must suppose then that the absorption of the quantum of ultra-violet light raises the electron to some level such as KL, from which it has a certain probability of falling to MN, where it will remain, provided that the incident light does not contain suitable frequencies which it can absorb

In the process described an electron was transferred to P from its next neighbour O. It may not, however, be necessary that the electron should come from an adjacent atom. The fact that under the usual applied voltages the formation of latent image is not accompanied by any photoelectric current that is measurable, merely shows that the electrons do not, on the average, travel for long distances. In conclusion it may be pointed out that if the ion P at the centre of the configuration is always a metal ion, the scheme will be in agreement with the neutral metal theory of the colouration, as advocated by Smekal.*

Summary.

In the quantum mechanical model of an insulating crystal there are bands of electron levels separated by wide zones of disallowed energies. In a perfect lattice these zones are empty and the occupied levels are not localized. Throughout an imperfect crystal are distributed numerous submicroscopic cracks, on the surfaces of which will be localized levels with energies lying in the empty zones. The formation of a latent image consists in the transfer of electrons to these localized energy levels in the disallowed zones.

---

* 'Phys. Z.,' vol 33, p. 204 (1932).

Boundary Conditions for the Wave Equation.

By W. H. McCrea and R. A. Newing.

(Communicated by S Chapman, F R S.—Received February 11, 1933.)

A single electron in the field of two fixed nuclei, constituting the idealized hydrogen molecular ion, provides the simplest case for the application of wave mechanics to molecular, as distinct from atomic, problems. The most extensive theoretical discussion of the corresponding wave equation has been given by A. H. Wilson* in these 'Proceedings.' He was led to conclude that this equation possesses no eigen-solutions satisfying the usual boundary conditions for an atomic problem Subsequent investigators† have succeeded, however, in obtaining by numerical methods eigen-values in good agreement with observed values of the energy But, with the exception of Teller, they appear not to have taken account of Wilson's result It is therefore worth while to investigate the existence of their solutions and to clear up, if possible, any doubt as to the applicability of the familiar boundary conditions to this type of problem.

The usual existence theorems‡ for eigen-values apply only to boundary conditions at ordinary points of the differential equation The difficulty in cases like Wilson's equation is that the conditions are given at singular points.

We have proved§ that, *if a, b are two successive singularities of the differential equation*

$$\frac{d}{dx}\left(\mathrm{K}\,\frac{dy}{dx}\right) - \mathrm{G}y = 0, \tag{1}$$

where K, G are functions of x and of a parameter λ, satisfying, in every closed interval (α, β), $a < \alpha \leqslant x \leqslant \beta < b$, the conditions usually imposed (Ince, *loc. cit.*) in the whole interval $a < x < b$, *and if the indicial equations for a, b have each the single non-negative root zero, then there exist real values of λ such that* (1) *possesses a solution which, with its derivatives, is finite and continuous throughout $a \leqslant x \leqslant b$.* We extend this to cases where only normal solutions exist at the end-points. Roughly speaking, it means that if equation (1) possesses one solution behaving in the required manner near $x = a$, and another solution behaving in the required manner near $x = b$, then it is possible so to choose λ that these solutions join up and become identical. The theorems themselves are of mainly mathematical interest, and will be published elsewhere.

---

* 'Proc. Roy. Soc.,' A, vol 118, p. 617 (1928)

† Morse and Stueckelberg, 'Phys Rev,' vol 33, p. 932 (1929); Teller, 'Z Physik,' vol 61, p. 458 (1930); Hylleraas, *ibid*, vol. 71, p. 736 (1931), where a fuller list of references is given.

‡ *Cf.* Ince, "Ordinary Differential Equations," chaps. 10, 11 (1926).

§ In a paper to be published by the London Mathematical Society.

This theory covers the important cases occurring in wave mechanics, and in particular it shows that Wilson's equation must possess eigen-solutions of the usual kind. These solutions seem to have evaded his search when he compares his general asymptotic forms of solution, on p. 625 of his paper, with a particular case (p. 624) in which certain coefficients do not vanish. This is sufficient to show that the corresponding coefficients in the general form do not vanish *identically*. But it does not follow that they cannot vanish, as required by the boundary conditions, for particular values (eigen-values) of the parameters, and in fact our work has shown that they must actually vanish in this way

---

The Adsorption of Iodine by Potassium Iodide.

By Brian Whipp, Hutchinson Research Student, St John's College, Cambridge.

(Laboratory of Colloid Science, Cambridge)

(Communicated by E. K Rideal, F.R S —Received February 16, 1933)

Opinion has been divided for some years as to whether the adsorption of a vapour on a solid surface is confined to a monomolecular layer until the vapour pressure reaches the saturation value, or whether multimolecular layers are formed.

Experimental work has generally failed to provide an unequivocal answer to this point, often because the true area of the surface of the adsorbent could not be estimated. It is now recognized that the true area of a surface may be many times the apparent area, and the ratio of these will depend on the previous treatment of the surface. The position has been complicated further by the difficulty of determining the amount of adsorption on a small area of a plane surface when the adsorption is so small as to be undetectable without sensitive apparatus ; consequently, the bulk of the experimental work has been carried out on finely divided adsorbents, of large surface. This breaking up of the adsorbent, whether accomplished by chemical or mechanical means, is bound to have an unpredictable effect on the ratio of the true to apparent area of the surface, as distinct from the purely geometrical multiplication. The use of porous or finely divided adsorbents may also introduce the additional complication of capillary condensation of vapours in the spaces between the particles. It is clear, as was first emphasized by Langmuir in 1918, that the investigation of adsorption on plane surfaces, using more delicate technique, is the first step towards a solution of the problem. Langmuir* measured the adsorption of gases on large plane surfaces of mica, glass and platinum at low temperatures and pressures. He showed that the adsorption reached a maximum, which corresponded to slightly less than one complete monomolecular layer, and that the extent of adsorption was related to the pressure by an equation of the form $x = \frac{abp}{1 + ap}$, where a and b are constants involving the condensation coefficient, the time of life of adsorbed molecules in the surface, and the number of "adsorbing centres."

Some evidence for the existence of a single layer only of molecules of a vapour is to be found in the work of Frazer, Patrick and Smith† and of Latham,‡ who showed that a freshly blown, "fire-polished" glass bulb adsorbed vapours to the extent of not more than one molecular layer at saturation pressure, but that after etching the glass with water or chromic acid, much larger quantities were taken up.

The determination of the adsorption of gases on crystals is important, not only because the crystals have approximately plane surfaces possessing a known number of adsorbing ions, but because, in simple cases at least, the attractive forces can be calculated. Thus Lennard-Jones and Dent§ found that the forces on an argon atom outside a potassium chloride crystal should become negligible at distances from the surface greater than the diameter of one atom.

A series of studies of the amounts of iodine adsorbed on condensed films of the halides of alkaline earth, has been made by de Boer.‖ Some of his earlier results were explained by an hypothesis which suggested as many as ten layers of adsorbed molecules but his later work* has yielded results which can only be explained on the theory of a monomolecular layer.

---

* 'J. Amer. Chem Soc.,' vol 40, p 1368 (1918).

† 'J. Phys. Chem.,' vol. 31, p. 897 (1927).

‡ 'J. Amer. Chem. Soc.,' vol. 50, p. 2987 (1928); *cf.* also Carver, *ibid.*, vol. 45, p. 63 (1923) and Curry, 'J. Phys. Chem.,' vol. 35, p. 859 (1931).

§ 'Trans. Faraday Soc.,' vol. 24, p. 100 (1928).

‖ 'Z. phys. Chem.,' B, vol. 13, p. 134 (1931).

The balance of evidence now available favours the theory of a monomolecular layer.† It should, however, be observed, (1) that uncertainties in determinations of the area may easily amount to 100 per cent., so that it is generally impossible to find more than the order of the number of layers adsorbed, and (2) that very few experiments throw light on the adsorption equilibria at saturation pressures.

There has, in fact, been no expression deduced which fits the experimental facts of adsorption of vapours over the complete range of pressure. Indeed, the results have generally been so unsatisfactory that the question of the existence of a definite adsorption maximum at saturation pressure has hardly been considered.

In the following paper sensitive adsorption measurements are described and definite evidence is presented for the existence of a bimolecular layer of iodine on potassium iodide crystals at saturation pressure, and an isotherm is suggested which fits the experimental results over the whole range of pressure. This expression contains only the two constants of Langmuir involving the extent of the adsorbing surface and the rates of evaporation of the molecules in the adsorbed phase.

Experimental.

The adsorbent was contained in the adsorption bulb A fig. 1. This was connected by a tap to a bulb B of known volume, acting as a charger, and to a Pirani pressure gauge P. Iodine was admitted by the tap T_4 to the charger, which was connected by the tap T_2 to a mercury condensation pump, McLeod gauge, a bulb of known volume C, mercury manometer M, and liquid air trap L. The charger was divided into two parts by the tap T_3 so that large or small charges of iodine could conveniently be admitted to the adsorbent.

The volumes of the various parts of the apparatus between the taps were found by admitting dry air into them from the standard calibrating bulb C. The pressures before and after expansion were read on the manometer M.

In an experiment, the apparatus was evacuated and the adsorbent baked out. The charger was filled with iodine vapour at a known pressure, and, after expanding this charge into the adsorption bulb, the pressure was read on the Pirani gauge The amount of adsorption was calculated from the difference between the calculated and observed pressures. Several expansions could be made and data for a complete isotherm obtained from one experiment.

---

* 'Z. phys. Chem.,' B, vol. 14, p. 457 (1931); B, vol. 15, p. 300 (1932); B, vol. 17, p. 161 (1932).

† But *cf.* Schlüter, 'Z. phys. Chem.,' A, vol. 153, p. 68 (1931).

The iodine was purified by three distillations from aqueous potassium iodide, which freed it from chlorine, bromine, iodine chloride and cyanogen iodide (the two latter being soluble in water). It was separated from the aqueous distillate, dried in a desiccator, placed in the tube t_3 and sealed to the apparatus. After being evacuated at 0° C. for about two hours on the condensation pump to free it from traces of water-vapour and gases, it was distilled *in vacuo* into t_2 and finally into t_1. The tubes t_3 and t_2 were then sealed off at S

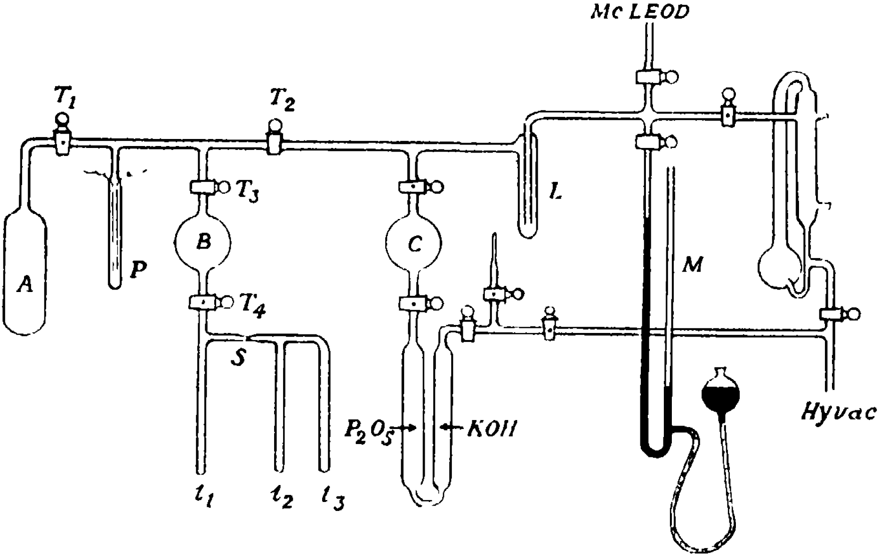


FIG. 1.

The Pirani gauge consisted of a platinum filament in a single loop about 10 cm. long, mounted on springy platinum-iridium supports, which kept it taut It was made of ribbon about 80 μ wide and 8 μ thick. In all experiments the filament was kept at a constant temperature by varying the current through it, so that its resistance, as indicated by a Wheatstone bridge arrangement, remained constant.* Pressures over the range $1 - 50 \times 10^{-3}$ mm. could be read with sufficient accuracy.

The Pirani gauge was immersed in a constant temperature bath at 20° C., and calibrated by maintaining various constant temperature mixtures round the iodine in i_1 and reading a voltmeter connected across the bridge when balanced. Eutectic mixtures were tried as constant temperature baths, but it was found more convenient to surround the iodine with a large unsilvered Dewar vessel full of HCl/ice freezing mixture. Although the temperature so obtained was not perfectly constant, it rose only slowly on account of the large heat capacity of the freezing mixture, and equilibrium vapour pressures were easily reached for each temperature of the bath. The temperature of the bath was read on a mercury thermometer, standardized at the National Physical Laboratory, and the pressure of the iodine corresponding to any temperature of the bath was obtained by interpolation from the data given in the International Critical Tables.

---

* Campbell, 'Proc. Phys. Soc.,' vol. 33, p. 287 (1921).

It was found that Shell Apiezon L grease was a good tap lubricant For a short time after greasing the taps, the grease evolved a gas when in contact with iodine (probably HI), but if the taps were well turned over a period of a few days while in contact with iodine vapour, the evolution ceased. A small absorption of iodine in the grease remained, but this could be made negligible (compared with the adsorption on the crystals) by using a large quantity of adsorbent and a large charging bulb.

By using a sensitive dead-beat pointer galvanometer across the Wheatstone bridge, and by controlling the heating current in the filament of the Pirani gauge by dial resistance boxes, the bridge may be balanced rapidly, and pressure readings taken within 1 and 1½ minutes of an expansion. Data for a complete isotherm could be obtained in about 12 to 15 minutes.

The disadvantage of using a large amount of adsorbent is that the time of baking out must be increased, and as it was found that the salt volatilized too quickly if a high temperature was employed, it was never baked at above 200° C. To obtain reproducible results the salt had to be kept at this temperature for about 20 hours in a high vacuum between experiments

Results

The temperature of the adsorbent was always 0° C.

A series of experiments was carried out with an empty adsorption vessel No instantaneous adsorption of iodine by the glass was noted and the data show that for iodine vapour deviations from Boyle's law are less than the experimental error, even up to saturation pressure.

The results of a typical experiment are shown in Table I.

Blank Experiment.

(Pressures are in mm. × 10^{-3}. Saturation pressure is 30 × 10^{-3} mm.)

Table I

Charge number	Observed pressure after expansion	Calculated pressure (assuming gas laws)
1	12·6	12·6
2	20·0	19·9
3	24·4	24·2
4	27·0	26·8
5	28·0	28·3

The maximum error is 1 per cent This represents the combined errors in the calibration of the apparatus and the readings of the Pirani gauge

The Isotherms

A series of experiments was carried out on the one sample of potassium iodide This consisted of about 115 gm of selected crystals of about 2 mm edge (B D H Analytical Reagent), and had an approximate area (estimated from the size of the crystals) of 1700 *cm* ^{2}. The volume of the dead space in the adsorption bulb was found by expanding into it dry air free from carbon dioxide from the charging system, and reading pressures on the manometer M Adsorption of air by the crystals was negligible.

When the amount adsorbed was plotted against the pressure, it was found .—

(1) That the lower pressure portion of the isotherm fitted the Langmuir equation $x = abp/(1 + ap)$ reasonably accurately

(2) That at higher pressures the adsorption increased almost linearly and tangentially to the original curve until saturation pressure was reached

(3) That the isotherms showed a definite maximum at the saturation pressure, when the amount of iodine adsorbed was, in all cases, roughly double that obtained by extrapolating to saturation pressure the Langmuir equation as determined from the initial portion of the isotherm.

(4) That the total number of molecules adsorbed at saturation was approximately twice the number of ions of one kind in the surface of the crystals used.

(5) **Apart from one or two irregularities, the amount adsorbed at saturation gradually decreased throughout the course of the exeriments.**

Some of the results are shown in Table II. Of the two Langmuir constants, *a* is proportional to the time of life of an adsorbed molecule in the surface, and *b* to the number of adsorbing centres. The adsorption is recorded as the difference between the observed and calculated pressures. These latter units may be reduced to gram moles at any time by multiplying by

$$\frac{410 \times 10^{-6}}{22{\cdot}4 \times 760} = 2{\cdot}40 \times 10^{-8},$$

410 c.c. being the free volume of the adsorbing bulb and charger.

The Langmuir constants in this table have been called "provisional" because they have been obtained from the first portion of the isotherms, that is, below a pressure of 4×10^{-3} mm In general (except in experiments which were not continued to higher pressures testing the application of the Langmuir equation to the initial portion) there was only one point on the isotherm below this pressure. The provisional constants were only used to calculate the ratio of the maximum adsorption calculated from the Langmuir equation to the total adsorption at saturation pressure. It will be seen that the mean ratio is 0·50(4) with a mean deviation from the mean of 10 per cent.

Table II

Experiment	Provisional Langmuir constants		Saturation adsorption moles $\times$ 2·40 $\times 10^{-8}$	Ratio Langmuir maximum/total
	b	*a*		
5	58·5	0·375	96	0·56
6	31·5	0·20	55	0·50
9	41·7	0·65	80	0·47
10	35·4	0·438	80	0·40(3)
12	39·2	0·403	70	0·51(2)
13	34·8	0·375	57·5	0·55(5)
16	24·8	0·274	40·5	0·54(8)
17	20·4	0·356	37	0·50(5)
19	23·4	0·312	41	0·41(5)
23	33·6	0·300	52·5	0·57(0)

Now as the Langmuir equation represents the formation of a single layer, ratios of about 0·5 must indicate the formation of a bi-molecular layer on the surface at saturation pressure. At low pressures the amount of iodine in the second layer will be negligible, but, with increasing pressure, the amount adsorbed will increase relative to that adsorbed in the first layer, until, at saturation, the two amounts are equal.

This prediction is confirmed by calculations on the area of the crystals covered. The distance apart of ions in the lattice of potassium iodide is 3·53 A. There are thus $4 \cdot 04 \times 10^{14}$ ions of each kind present per square centimetre of surface. Now in the (1, 1, 1) plane of an iodine crystal there are $4 \cdot 35 \times 10^{14}$ iodine molecules per cm.^{2}* so that it is reasonable to assume that the iodine molecules will only be adsorbed on ions of one kind. If they were adsorbed on both kinds of ions, impossibly close packing would result.

The simple geometrical area of the crystals has been estimated by ordinary mensuration as approximately 1700 cm ^{2} on the assumption that they are 2 mm. cubes. It is pointless to attempt to estimate the area of the surface with greater accuracy. Thus the total number of adsorbing ions is

$$1700 \times 4 \cdot 04 \times 10^{14} = 6 \cdot 9 \times 10^{17}.$$

The amount of iodine adsorbed at saturation decreases throughout the experiments. This shows that the baking or the treatment with iodine is decreasing the effective area of the adsorbent. If the maximum amount of adsorption recorded is divided by two and expressed as a number of molecules, we obtain $6 \cdot 8 \times 10^{17}$ for the first layer.

It happens that the agreement is even better than would have sufficed or been expected for the proof of the theory.

The Kinetics of Adsorption

The most logical interpretation of the facts is to assume the existence of a strongly adsorbed layer of molecules attracting to itself a second layer by the action of much weaker fields. The forces binding the molecules adsorbed in the second layer will be much less than those binding the molecules in the first, so that the rate of evaporation of the former molecules will be higher than that of the latter.

The number of adsorbing centres for the second layer must be the same as for the first because the amounts adsorbed in the two layers are equal at saturation pressure.

If it be assumed that iodine is adsorbed in the second layer according to simple Langmuir kinetics represented by the equation $x = a'bp/(1 + a'p)$, where b is proportional to the amount of the surface covered by the first layer, while a' is to a, of the original Langmuir equation, as the rate of evaporation of molecules in the first layer is to that of those in the second—and is so small that direct proportionality is given up to saturation pressure—then the final gradient of the adsorption curve, as calculated from the equation, is ten times too small. It is thus clear that the second layer cannot be built up by a mechanism similar to that obtaining for the first layer, and that some other process is involved.

---

* Langmuir, 'J. Amer. Chem. Soc.,' vol. 54, p. 2812 (1932).

The Mechanism of Formation of the Second Layer.

In this section an attempt is made to derive an equation which shall fit the experimental results.

Let ν be the rate of evaporation from unit area of first layer when completely formed (assuming no second layer present to impede evaporation), and let ν' be the corresponding quantity for the second layer. Let θ and θ' be the fractions of total surface covered with the first and second layers respectively.

Let μ be the number of gram moles striking 1 cm.^{2} of surface per second, then

$$\mu = \frac{p}{\sqrt{2\pi \mathrm{MRT}}}, \tag{1.0}$$

where p is the pressure and M the molecular weight of the gas.

The rate of condensation of molecules on the first layer is $\alpha\mu\,(1-\theta)$ where α is the condensation coefficient. The rate of evaporation is given by $\nu\,(\theta-\theta')$ on the assumption that evaporation from the first layer cannot take place when this is covered by a second layer. Hence, for equilibrium of the first layer,

$$\alpha\mu\,(1-\theta) = \nu\,(\theta-\theta'). \tag{1.1}$$

Similarly for the second layer the rate of evaporation is $\nu'\theta'$. If we assume that condensation occurs in the second layer whenever a molecule from the gas phase strikes the first layer, whether or no the first layer is already covered by a second layer, the rate of condensation will be $\alpha'\mu\theta$, where α' is the condensation coefficient for the second layer. This implies that the molecules in the second layer are so loosely held that they are pushed aside laterally by the impact of fresh molecules from the gas phase, until the first layer is completely covered.

Condensation into the second layer is thus similar to condensation of iodine vapour on to a solid surface of iodine and involves the assumption of a low critical energy increment for lateral mobility.

Equating the rates of condensation and evaporation for the second layer,

$$\alpha'\mu\theta = \nu'\theta'. \tag{1.2}$$

Substituting the value of θ from (1.2) in (1.1)

$$\alpha\mu\left\{1 - \frac{\theta'\nu'}{\alpha'\mu}\right\} = \nu\theta'\left\{\frac{\nu'}{\alpha'\mu} - 1\right\}. \tag{1.3}$$

Now

$$\frac{\alpha\mu}{\nu} = ap \quad \text{and} \quad \frac{\alpha'\mu}{\nu'} = a'p, \tag{1.4}$$

where p is the pressure in the gas phase.

The only properties of the surface upon which a and a' are dependent are the condensation coefficients, the number of adsorbing ions per unit area and the time of life of adsorbed molecules in the surface so that they may be considered as the respective mean lives of the molecules adsorbed in the first and second layers.

Hence

$$ap\left\{1 - \frac{\theta'}{a'p}\right\} = \theta'\left\{\frac{1}{a'p} - 1\right\} \tag{1.5}$$

or

$$\theta' = \frac{a\,a'p^2}{1 + p(a - a')}. \tag{1.6}$$

Similarly,

$$\theta = ap/\{1 + p\,(a - a')\}. \tag{1.7}$$

If x_1 be the amount of iodine adsorbed in the first layer and x_2 that adsorbed in the second, then

$$x_1 = \frac{N'\theta}{N} \quad \text{and} \quad x_2 = \frac{N'\theta'}{N},$$

where N′ is the number of adsorbing spaces and N is the Avogadro number.

Therefore from (1.6) and (1.7)

$$x_1 = \frac{abp}{1 + p\,(a - a')}. \tag{1.8}$$

and

$$x_2 = \frac{a\,a'bp^2}{1 + p\,(a - a')}, \tag{1.9}$$

where b ($= N'/N$) is the Langmuir constant representing the number of adsorbing centres.

Thus, X, the total adsorption ($= x_1 + x_2$), is given by

$$X = \frac{abp\,(1 + a'p)}{1 + p\,(a - a')} . \tag{2.0}$$

At the saturation pressure $x_1 = x_2$. Equating x_1 and x_2 and writing p_0 for the saturation pressure, we find

$$a'p_0 = 1. \tag{2.1}$$

Thus a' which is proportional to the mean life of molecules in the second layer is not an arbitrary constant, but is simply related to the saturation vapour pressure.

Substituting the value of a' from (2.1) in (2.0), we obtain finally,

$$X = \frac{abp\,(1 + p/p_0)}{1 + ap - p/p_0} \tag{2.2}$$

as the adsorption isotherm.

At the saturation pressure $X = 2b$ as would be expected.

If we consider the kinetics of evaporation and condensation of molecules of iodine vapour on a solid iodine surface, we find that for equilibrium,

$$\alpha''\mu = \nu'',$$

where α'' is the condensation coefficient and ν'' is the rate of evaporation for a solid iodine surface.

Thus assuming the same value of the condensation coefficients α', α'',

$$a''p_0 = 1,$$

where a'' represents the mean life of the molecules in the surface layer of a solid iodine crystal. Thus (by (2.1)) the rate of evaporation of molecules in the second layer is the same as that of molecules on a pure iodine surface, provided the coefficients α', α'' are equal.

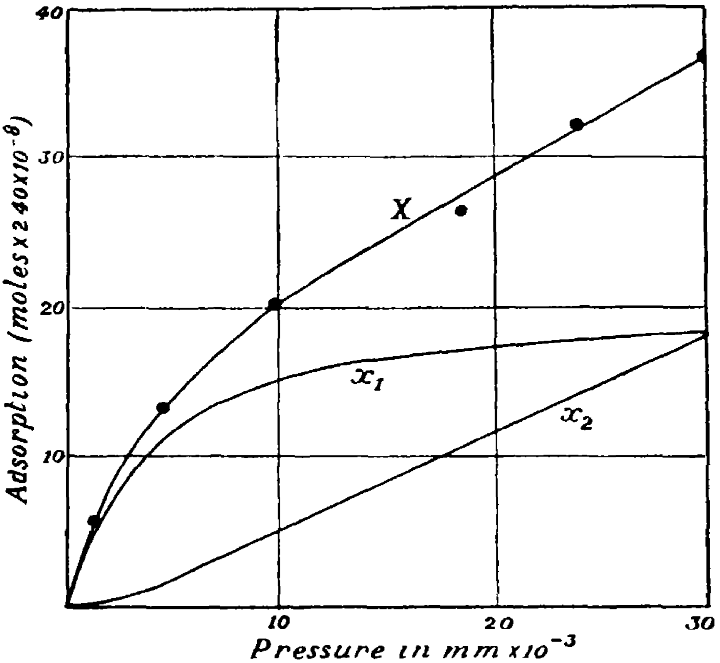
In Table III are the values of X calculated from the isotherm (2.2) together with the amounts actually observed (interpolated at the pressures for which X has been calculated). In each space in the table the calculated amount is placed above the observed amount. Underlined figures show where the theoretical value was equated to the experimental value for the purpose of calculating the constants a and b. At the bottom of the table the values of the constants for the several isotherms have been tabulated.

Table III.

Isotherm.	1.	2.	4	6.	16.	17.
Pressure— 2×10^{-3} mm.	21·4 21·2	22·7 22·7	27·9 29·7	8·4 8·8	8·5 9·0	8·1 8·2
4×10^{-3} mm	32·1 32·0	30·9 30·9	38·7 38·7	14·0 14·0	13·0 13·0	12·5 12·5
6×10^{-3} mm	39·3 39·8	36·2 36·5	45·7 45·2	18·25 18·0	— —	— —
8×10^{-3} mm.	44·9 45·2	40·5 40·2	51·0 50·4	21·7 21·2	18·7 18·5	18·2 18·0
10×10^{-3} mm.	49·7 50·0	43·8 44·3	55·7 55·5	— —	— —	— —
12×10^{-3} mm.	54·0 54·0	47·0 47·0	60·0 60·0	27·6 26·5	22·7 22·3	22·4 22·1
14×10^{-3} mm.	58·2 57·8	50·0 49·7	64·2 64·4	— —	— —	— —
16×10^{-3} mm	62·2 61·0	52·9 51·0	68·2 69·7	32·6 32·0	26·2 25·5	25·7 25·1
20×10^{-3} mm.	— —	— —	— —	37·5 37·5	— —	— —
22×10^{-3} mm	— —	— —	— —	— —	31·0 31·0	— —
25×10^{-3} mm.	— —	— —	— —	42·8 45·5	33·4 35·0	32·8 32·5
28×10^{-3} mm.	— —	— —	— —	46·0 50·0	— —	35·0 35·0
30×10^{-3} mm	— —	— —	— —	48·4 52·5	37·4 40·6	37·0 37·3
Langmuir constants— (*a*) (*b*)	0·400 43·6	0·665 36·3	0·590 46·7	0·259 22·7	0·346 18·65	0·325 18·45

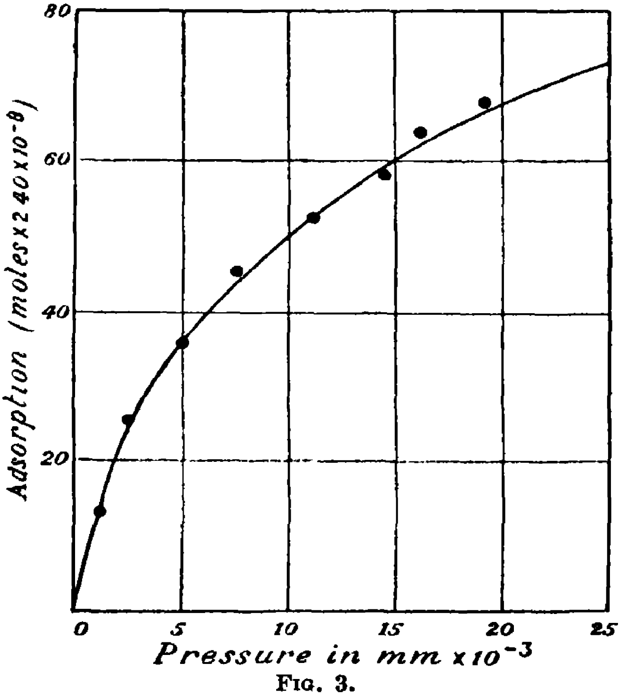
The amounts adsorbed are in units of moles × 2·40 × 10^{-6}, the calculated values being inserted above the observed values.

In isotherm 1, fig. 3, the data are shown up to about two-thirds saturation pressure whilst in fig. 2 the complete analysis of isotherm 17 is shown up to actual saturation, x_1 and x_2 being the amounts adsorbed in the first and second layers respectively (calculated from equations (1.8) and (1.9)). In each case the experimental data (not interpolated) have been plotted as points, and the



40
30
20
10
Adsorption (moles x 2 40 x 10⁻⁸)
X
x₁
x₂
10
20
30
Pressure in mm x 10⁻³

FIG. 2.



80
60
40
20
Adsorption (moles x 2 40 x 10⁻⁸)
0
5
10
15
20
25
Pressure in mm x 10⁻³

FIG. 3.

values calculated from equations (2.2), (1.8) and (1.9) as the smooth curves. It will be seen that there is good agreement between the experimental and theoretical values.

The value of a' deduced from equation (2.1) is about ten times smaller than that of a, and so at low pressures its effect is negligible and the isotherm reduces to a simple Langmuir expression.

In his original work of 1918, Langmuir (*loc. cit.*) found the constants a and b of his simple equation to vary irregularly throughout the course of the experiments and the same effect has been observed in the present work. This is in all probability owing to an alteration in the surface of the adsorbent caused by the long process of baking-out.

When taking the data for isotherms 6 and 16 at pressures near to saturation, the iodine pressure in the apparatus fell slowly after each expansion. This was shown to be caused by absorption of iodine by the grease, owing to the spreading of the grease from the taps during a period of hot weather. In isotherm 17 the pressures were read as soon as possible after each expansion and the curve is seen to fit the theoretical expression up to saturation pressure.

Discussion.

Langmuir* has shown from the data given by Harris, Mack and Blake† that adsorbed films of iodine are not likely to contain more than $4 \cdot 35 \times 10^{14}$ molecules of iodine per cm.^{2} (the number per cm.^{2} in the (1, 1, 1) face of an iodine crystal) after the first layer because repulsive forces will become operative at greater densities of packing. On these grounds he criticized the earlier theory of de Boer (*loc. cit.*), which demanded vertical rows of iodine atoms above each anion of salt, giving a layer density of $7 \cdot 75 \times 10^{14}$ molecules/cm.^{2}. Langmuir considered the arrangement of molecules in the adsorbed film to be generally quite irregular after the first layer.

The potassium iodide crystal appears to be peculiarly suited to act as a foundation of stable layers of iodine molecules because the number of ions of one kind per cm.^{2} is $4 \cdot 04 \times 10^{14}$. Iodine molecules being adsorbed on ions of one kind, and the distance between K ions and I ions being 3·52 A., it follows that the smallest distance apart of the adsorbed molecules is $\sqrt{2 \times (3 \cdot 52)^2}$ A. and the next smallest $2 \times 3 \cdot 52$ A.

These distances are approximately the same as those found by Harris, Mack and Blake (*loc. cit.*) for two of the sides of the unit cell of iodine. It may therefore be expected that the forces between the iodine molecules in a bi-molecular layer adsorbed on potassium iodide are of the same order as the homo-polar forces between the molecules in a pure iodine crystal, and thus the bi-molecular layer attains equilibrium with the vapour at the saturation pressure of pure iodine. Optical examination of the adsorbed films designed to show (*a*) the thickness of the layer as in the experiments of Frazer,* and (*b*) the polarization of the iodine by measurement of its absorption spectra, as in the work of Chilton and Rabinowitsch† would be of interest in this connection.

---

* 'J. Amer. Chem. Soc.,' vol. 54, p. 2812 (1932).

† 'J. Amer. Chem. Soc.,' vol. 50, p.1583 (1932).

Chilton and Rabinowitsch show that the absorption spectra of iodine adsorbed on calcium fluoride (as measured by de Boer) and on chasabite are similar to that of solid iodine. The adsorbed films in this work were of a yellowish-brown colour so that it would seem from the colours of the films observed by Chilton and Rabinowitsch that the molecules are polarized to the same extent as those in solid iodine

We have assumed that there is lateral diffusion in the second adsorbed layer and if this layer is similar to the surface layer of an iodine crystal, as we have indicated, we should anticipate the phenomenon of surface diffusion in the latter case. This prediction is of interest in connection with the respective theories of crystal growth of Kossel and Stranski on the one hand, and of Volmer on the other.

According to Kossel‡ the limiting factor in the growth of crystals is the slow rate at which a lattice plane is built up. Kossel and Stranski§ imagine that the molecules are regularly orientated on the surface and leave out of account the possibility of lateral mobility. Volmer‖ on the other hand concludes that the condensation of isolated molecules and small aggregates is always a much more rapid process than that deduced by Kossel and Stranski, and regards the surface as covered by a two-dimensional mobile layer; the velocity of crystallization being governed by the rate at which lattice nuclei rea formed by lateral diffusion. A nucleus of given size is only in equilibrium with the two-dimensional mobile layer at a definite degree of supersaturation, the degree of supersaturation necessary increasing as the size of the nucleus becomes smaller, so that as the degree of supersaturation is reduced, the rate of growth of the nucleus will fall, reaching zero before the equilibrium (saturation) pressure is reached. This region of zero growth of a crystal in its still supersaturated vapour has been observed directly by Volmer (*loc. cit.*) for iodine, thus supporting the existence of the mobile surface layer.

---

* 'Phys. Rev.,' vol. 34, p. 644 (1929).

† 'Z. phys. Chem.,' B, vol. 19, p. 110 (1932).

‡ 'Nachr. Ges. Wiss. Göttingen,' p. 135 (1927).

§ 'Z. phys. Chem.,' A, vol. 136, p. 259 (1928).

‖ 'Z. phys. Chem.,' A, vol. 156, p. 1 (1931).

The fact that the second adsorbed layer when full contains the same number of molecules as the first seems to indicate that, for iodine on potassium iodide at least, lateral mobility is subject to certain constraints by the hetero-polar substrate.

It may be noted that the existence of KI_3 as a separate phase was disproved. Experiments were carried out in which the pressure of iodine above a bimolecular layer of adsorbed iodine was slowly reduced by absorption of the iodine vapour in tap grease. These were continued until only a fraction of a single layer remained adsorbed, but in no experiment did the pressure show any stationary value below that of saturation. In a review of the evidence in favour of the existence of potassium tri-iodide as a solid phase, Bancroft, Scherer and Gould* show that it is extremely unlikely that solid KI_3 has been prepared, although this does not exclude the possibility of its existence in solution.

My great thanks are due to Professor E. K. Rideal, F.R.S., for suggesting this work and for his valuable encouragement and criticism, and to Dr. O. H. Wansbrough-Jones for much helpful advice.

I also wish to express my thanks to the Goldsmiths' Company and to the Department of Scientific and Industrial Research for grants which enabled the investigation to be carried out.

Summary.

Adsorbed films of iodine on potassium iodide crystals are found to be two molecules thick at the saturation pressure. A kinetic explanation of the process is advanced on the assumption of a loosely bound second layer, and a two-constant isotherm is deduced which represents the adsorption up to saturation pressure. Evidence is presented showing that the surface molecules in the bi-molecular film have properties similar to those in the surface of a crystal of pure iodine and the bearing of this on Volmer's theory of crystal growth is considered. Potassium tri-iodide is not formed as a separate phase.

---

* 'J. phys. Chem.,' vol. 35, p. 764 (1931).

The Observation of Gravity by Means of Invariable Pendulums.

By E. C. Bullard, Department of Geodesy and Geophysics, Cambridge.

(Communicated by Sir Gerald P. Lenox-Conyngham F R.S —Received March 17, 1933)

[Plate 3.]

1. *Introduction.*

By observing the variation of the gravitational field over the earth's surface it is possible to obtain information about the distribution of attracting matter within the earth. Of the methods of investigating these variations, the observation of the change in period of invariable pendulums carried from place to place has so far yielded most information. Until recently such observations have been of necessity extremely lengthy, and only about 3 per cent. of the earth's surface has been at all adequately surveyed.* It is therefore important to develop rapid methods of observing, and recently several papers† have appeared describing methods of using wireless signals in timing pendulums. The purpose of the present paper is to describe the methods developed in Cambridge. They differ from former methods in that :—

(1) No special signals are required.
(2) The apparatus is considerably more portable than any so far described.
(3) The errors in timing are so small that, with a 1-hour swing, they are negligible compared with other sources of uncertainty.

2. *Accuracy Required.*

The accuracy to which it is profitable to observe gravity depends on the object of the investigation undertaken. For most geophysical purposes, such as discussions of isostasy, the variations produced by very local causes cannot be taken into account, as the detailed geological information necessary is not available, and in any case the labour involved would be prohibitive ; further, the accuracy with which the height and latitude of a station can be fixed in undeveloped country is limited. Now a change of 0·001 cm./sec.^{2} in the intensity of gravity is produced by :—

---

* J. de Graaff Hunter, 'Mon. Not. R. Astr. Soc., Geophys. Suppl.,' vol. 3, p. 42 (1932).

† Andersen, 'Ver. Balt. Geod. Komm.,' p. 215 (1931) ; Meisser, Martin, and Gengler, 'Beitr. Geophys.,' Suppl. vol. 2, p. 131 (1932).

(*a*) 3 metres change of height.
(*b*) 1·3 km. change of latitude (at lat. 52°).
(*c*) The attraction of a layer of average rock 10 metres thick.

Thus it may be said that in most cases no use can be made of an accuracy greater than 1 or 2 parts in a million. For special purposes, however, such as prospecting for oil, an accuracy 5 to 10 times as great may be required, but beyond this any further increase would seem quite useless.*

The problem is therefore, to develop a method of timing pendulums which, shall give g to an accuracy of the order of 1 in a million in as short a time with as portable an apparatus as possible.

3. *General Description of Methods.*

Two methods of working have been used :—

(1) Records of two wireless signals controlled by an observatory are taken, the pendulums being allowed to swing continuously during the interval between them. This method, although simple and accurate, has several disadvantages :

(*a*) Owing to unexplained systematic differences between the corrections to the same time signals given by different observatories it is desirable that the initial and final signals should originate from the same observatory. The most convenient pair of signals are those emitted from Rugby at 10 and 18 hours, and even these are so far apart that only one observation can be made at a station in a day.
(*b*) The accuracy depends on the accuracy of observations at the observatory, over which the gravity observer has, as a rule, no control, and of which it is difficult for him to estimate the accuracy.

The method has, however, been used with success.

(2) An arbitrary morse message transmitted by a high powered station may be recorded along with the swings of the pendulums at the field station; while an exactly similar record is made simultaneously of another pendulum swinging at a fixed base station. After a predetermined time (in practice

---

* This does not, of course, apply to measurements of the variation of gravity with time at a given place, where the above uncertainties do not exist. The methods employed for such measurements are, however, entirely different (Tomashek, 'Astr. Nach.,' vol. 244, p. 257 (1931)).

1 hour) a further similar pair of records is taken. Corresponding places in the morse messages on the field and base station records then represent simultaneous instants, and it is possible to determine by measurements on a pair of records what is the difference of phase between the field and base station pendulums at any instant. By determining this phase difference at several points on the initial pair of records it is possible to find approximately how fast the field pendulum is gaining or losing on the base station pendulum; this suffices to determine the whole number of swings gained or lost during the hour. Then by measurements on both initial and final pairs the fractional part of a swing gained or lost may be found, and, if the total number of swings at the base station be known, the difference of period may then be determined. From the difference in this quantity when one set of pendulums is swung at the field station and one at the base station, and when both are swung side by side at the base station, the change in period of the pendulums on going from the base to the field station may be found, and thence the difference in g; provided the period of the base station pendulum be approximately known.

No difficulty has been experienced in practice in starting the two chronographs sufficiently nearly simultaneously. Both observers turn on their wireless receivers at the pre-arranged time (given by their wrist watches), and wait for a signal. As soon as a signal arrives the chronographs are started and left to run for about 5 minutes. No record has ever been lost through the observers starting their chronographs at different times or through insufficient signal being recorded. Alternatively, signals (such as weather reports, news bulletins, and time signals) which occur at definite times may be used.

4. *Detailed Description of Apparatus.*

(*a*) *Pendulum Cases.*—The cases were designed by Sir Gerald Lenox-Conyngham, and made by the Cambridge Instrument Company. That used at the field station consisted of an "Alpax"* box which can be evacuated, and which is provided with the usual arrangements for lowering the two pendulums on to the agate blocks on which they swing, and for starting them simultaneously with equal amplitudes and opposite phases. The pendulum at the base station is swung in the front compartment of the three-pendulum case which has been described elsewhere.† The pendulums have also already been described.† They are of the von Sterneck form, and each is machined from a single forging of "Nilex" (a form of invar), the hard stellite knife-edges being afterwards fastened to the pendulum head by wedges.

---

* Aluminium with 8 to 13 per cent. of silicon.

† Lenox-Conyngham, 'Geog. J.,' vol. 73, p. 326 (1929).

The boxes are lined with 0·25 mm. of Mumetal to eliminate changes of period due to variations in the earth's magnetic field and in the magnetic moment of the pendulums.*

(*b*) *Chronograph.*—A photographic method of recording the wireless signals and the motion of the pendulums was chosen, as it requires less subsidiary apparatus than alternative methods, and there is little to get out of order.

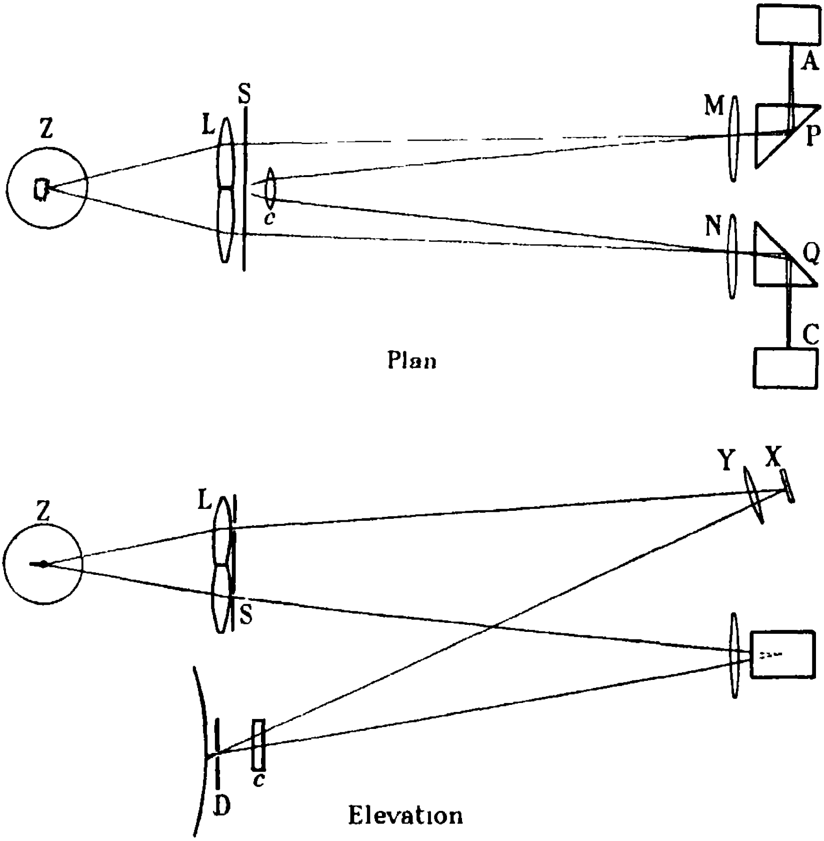


FIG. 1.—Optical system, not to scale.

The optical system employed is indicated in fig. 1. Light from a 6-volt 24-watt motor car headlamp bulb Z, fed from an accumulator, is divided into three beams by a split lens L. Two of these beams are focussed by L on the two pendulum mirrors A and C, passing through the totally reflecting prisms P and Q, which turn them through a right angle, and through the lenses M and N. The third beam is focussed on the oscillograph mirror X. The light incident on the pendulum mirrors retraces its path through the prisms and passes a second time through the lenses M and N, which form images of two ½ mm. horizontal slits S placed in front of L. These images are concentrated in a horizontal plane by a cylindrical lens c, with its axis vertical. The positions of these images when the pendulums are at rest are adjusted by rotating the prisms P and Q so that they fall on a horizontal slit D. Thus as the pendulums swing the images move up and down, and at each passage through the equilibrium position a flash of light passes through the slit D. Immediately behind this slit is a drum carrying a sheet of photographic paper. This drum is rotated about a horizontal axis by a spring-driven gramophone motor, and at the same time traversed slowly parallel to its axis. The axle A, fig. 2, driven by the motor is threaded for half its length. One end E of the drum D slides freely on the unthreaded half of the axle, the other end, F, is carried by the outer part of a ball bearing B. The inner part of this bearing carries a threaded bush C working on the threaded part of the axle. This bush is prevented from turning by a bar X fixed to it, a slot in one end of the bar slides on a rod R parallel to the axle. The drum is turned by a bar G sliding freely in a hole in the end plate E of the drum. The lower portion of the bush C is cut away, so that if the drum is slightly lifted it can be disengaged from the thread and slid back to its starting-point. This arrangement has the advantage over the usual one, in which the axle and drum are traversed together, in that the overall length is only twice that of the drum instead of three times.

---

* A paper on the effect of a magnetic field on the period of an invar pendulum and its elimination is to appear shortly in the 'Proc. Camb. Phil. Soc.,' vol. 29 (1933).

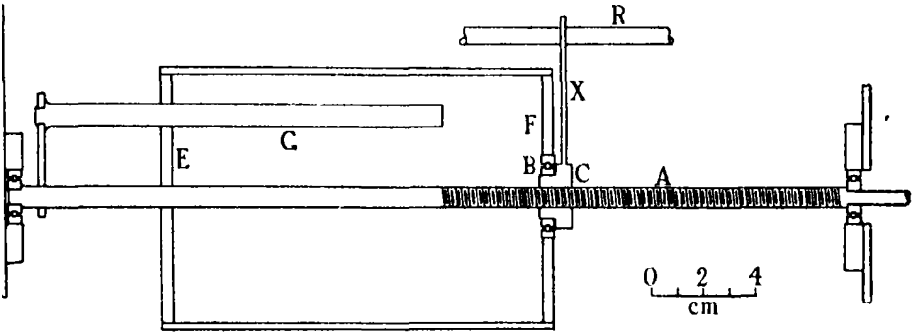


FIG. 2.—Chronograph drum.

The drum is 10 cm. in diameter, and 15 cm. long, is rotated once in 5 seconds, and moved along 2·5 mm. per revolution. The peripheral velocity is therefore about 7 cm./sec., and the record is a spiral having 60 turns.

As it was found that small displacements of the prisms P and Q, fig. 1, were liable to occur during evacuation, small vertical adjustments were provided for the slits S. An adjustment for the height of the lamp bulb Z, is also provided, so that the image of its filament formed by L can be brought on to the pendulum mirrors. As the lens L is not moved these two adjustments are independent, and altering one does not affect the other.

The chronograph is mounted in a wooden box, the lower half of which contains the motor and drum, and a light-tight compartment for storing photographic paper. The paper is put on the drum by inserting the hands through a black velvet bag. The top half of the box contains the light, lens, and adjustable slit system. It also serves as space for carrying the oscillograph, levels, and other loose parts.

The light provided by the bulb and accumulator is amply sufficient, and it would probably not be difficult to work with a smaller bulb and dry batteries, should the accumulator be inconvenient.

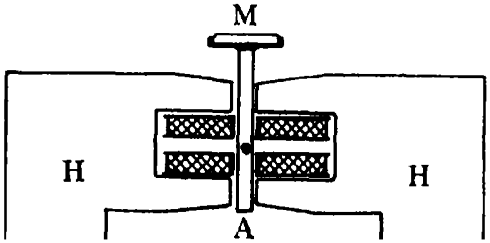


Fig. 3 —Section through oscillograph.

(*c*) *Oscillograph.*—The oscillograph is of the type described by Wynne-Williams,* and is a converted "balanced armature" loudspeaker. The armature, A, fig. 3, which carries a mirror M, is magnetized by the signal currents from the wireless set, and is rotated by the attraction and repulsion of a permanent magnet H. The arrangement forms a very robust but somewhat insensitive oscillograph. It is probable that a specially designed instrument would be better.

---

* 'Proc. Roy. Soc.,' A, vol. 131, p. 391 (1931).

The oscillograph is placed in a magnetically shielded box on top of the pendulum case, and receives the third beam of light from the split lens L, fig. 1. An image of a third slit in front of L is formed just below the slit of the chronograph by a lens Y in front of the oscillograph. When the oscillograph is supplied with a heterodyned continuous wave signal the spot of light vibrates with the frequency of the heterodyne note (about 900 cycles/sec is the frequency ordinarily employed), and crosses the slit twice in each vibration. Owing to the width of the slits the marks produced fuse together, giving a black bar on the record. Thus a morse message is recorded as a spiral of black marks in which a faint 900 cycle structure can be seen under the microscope. A portion of such a record is shown in fig. 9, Plate 3.

(*d*) *Wireless Receiver*.—The receiver is of orthodox design employing a frame aerial, one stage of screened grid high frequency amplification, coupled by a tuned plate circuit (whose tuning is mechanically linked to that of the aerial circuit) to a leaky grid rectifier, followed by two stages of transformer-coupled low frequency amplification. The first low frequency stage has a parallel resonance circuit between its grid and filament tuned to the resonant frequency of the oscillograph. This greatly improves the selectivity of the system, it also renders it more stable, as the condenser of the resonant circuit helps to shunt high frequency currents to earth. The set is provided with a separate oscillator, owing to the sharp audio-frequency resonance curve of the amplifier the tuning of this circuit is extremely critical, and special care has been taken to proportion the circuits so that changes of battery voltage do not produce changes of frequency.* The aerial, batteries, and amplifier, are all contained in the same box. The wave band covered is 15000—20000 metres.

The pendulum case, chronograph, and wireless set described are all that is required for a gravity determination when time signals controlled by an observatory are used.

For the second method described above, using an arbitrary morse signal, the apparatus at the field station is the same as for the method using time signals. At the base station the apparatus is the same as at the field station except that only one pendulum is swung. In addition it is necessary, for reasons discussed below, to provide 1/100 second marks on the record at the base station. A vibrator driven from the 50 cycle alternating current mains interrupts a beam of light twice in every period of the supply. This beam of light is arranged to fall on the slit of the chronograph, and gives a spiral of marks, the morse message and the pendulum marks lying on either side of it. A portion of a record is shown in fig. 9, *c*, Plate 3. It is further necessary to count the number of swings of the base station pendulum. An arrangement employing a photocell would probably be quite satisfactory,* but it happened that an old astronomical clock was available and provided a simpler solution. Its pendulum was lengthened so that it had a period of 1·0108 seconds, twice that of the base station pendulum. Electrical impulses from this clock operated a telephone message counter, to the ten's wheel of which an arm had been attached. Once in each revolution (*i.e.*, once in every 200 swings of the gravity pendulum), this arm cut off the beam of light marking 1/100 second for about 1/50 of a second. The resulting gaps in the spiral of 1/100 second marks provide a means of numbering the pendulum swings. As the swings are only carried on for an hour, no great accuracy is necessary in adjusting the clock; all that is necessary is that it should not be more than a few tenths of a swing out in the hour. In any case each record contains two or three marks, and the mutual consistency of these provides a satisfactory check.

---

* F. B. Llewellyn, 'Proc. Inst. Radio Eng.,' vol. 19, p. 2063 (1931).

5 *Methods of Reduction.*

When using time signals the reduction is very simple. The 305 dots of the rhythmic signal appears as 305 bars, each about 1 cm long, every 61st being lengthened to 3 cm. The pendulum swings appears as pairs of dots on each side of these, the pairs being alternately sharp and diffuse as the light moves alternately with and against the direction of motion of the recording paper. A portion of such a record is shown in fig. 9, *a*, Plate 3. About every 18 seconds a "coincidence" occurs between the beginning of a time signal dot and a pendulum mark. In general the coincidence is not exact, one pendulum mark falling before the beginning of one time signal mark, and the next falling after the next time signal mark. The fraction of a time signal interval at which the "true coincidence" occurs may be estimated by eye, or may be found more accurately by measuring, with a transparent scale, the distances between the pendulum marks and the time marks for several swings on each side of coincidence, plotting against the number of the time mark and reading off the intercept. A typical example of such a graph is given in fig. 4. The uncertainty judged from the scatter of the points from a straight line is less than 0·1 time signal intervals; as the pendulums take about 37 time signal intervals to lose a complete swing on the time signal, this corresponds to about 1/370 second of time. Each record contains about 8 coincidences of each pendulum when moving up, and an equal number when moving down, the two records therefore contain 64 coincidences altogether. From these an approximate value of the coincidence period can be derived, which can be used to obtain the whole number of coincidences during the 8 hours between the signals. This number is about 800, and the accuracy of measurement appears sufficient to determine it unambiguously. In practice, however, the number of coincidences calculated from the approximate coincidence period is not as near to an integer as would be expected. This is probably due to the action of the "hit-and-miss synchronizer" on the Shortt clock controlling the Rugby signals; this may not act at all during the signal, or may act once, or twice, introducing a discontinuity of about 1/300th second each time.* All uncertainty on this account can be eliminated by taking an intermediate signal (say, the Nauen 12-hour signal), and using it and the 10-hour Rugby signal to obtain a closer approximation to the coincidence period, and then using this to determine the whole number of coincidences in 8 hours. The

---

* A "phase controlled" thyratron circuit would provide a very simple amplifier. Hull, 'Gen. Elect. Rev.,' vol. 32, pp. 213, 390 (1929).

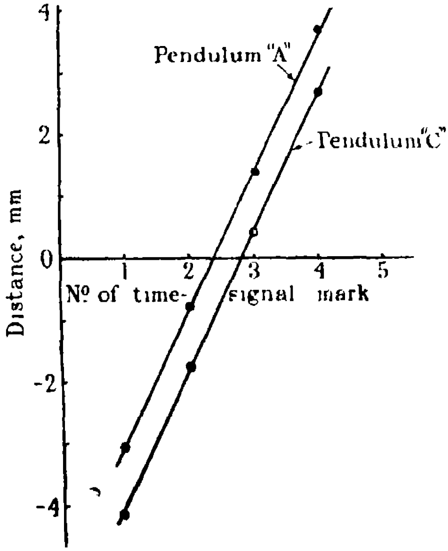


FIG. 4.—Graph for determining instant of coincidence.

---

* I am indebted to Dr. Jackson of Greenwich Observatory for this information.

differences in the dot numbers for coincidences on the morning and evening signals are then found, and the mean taken. The deviations of these from the mean indicate a probable error of 0·07 dot interval in the determination of a single coincidence, which corresponds to 1/500th second or 1 part in 15 million on the 8 hours. As we have 16 values of this difference for each pendulum, the accuracy of this part of the observation is considerably greater than is necessary, and a swing of 1 or 2 hours would produce all that is required, if signals of the necessary reliability were available. From this mean and the known number of coincidences a final value of the coincidence period may be found, and the time of swing calculated.

Averaging the coincidences with the spot of light moving up and with it moving down eliminates errors due to the position of the spot when the pendulums are at rest not exactly coinciding with the slit S, and also errors due to small movements of the slits or prisms during the swing. Changes in the rate of revolution of the drum can only affect the fractional part of the dot number at which a coincidence occurs, and are entirely negligible in practice.

From observations at a number of stations in various parts of England the probable error of a single observation, deduced from the consistency of the results of the individual swings made on a single visit to a station is $1 \cdot 0 \times 10^{-7}$ second on the mean period of the two pendulums, or 0·0004 cm./sec.^{2} on g. This includes uncertainties in the time signal corrections, but not errors produced by changes in length of the pendulums between stations.

(*b*) *Method Using an Arbitrary Signal.*—Eight points on the morse message on the field station record are marked. These are chosen so that :—

(1) A pair of pendulum marks is as near as possible to the chosen morse mark, thus reducing errors due to irregularities in the motion of the recording paper.

(2) Four of the marks are near pendulum marks made by the light moving in the same direction as the paper (sharp marks), and four near to those made by the light moving in the opposite direction (diffuse marks). Averaging results obtained from the two sets eliminates errors due to the chronograph slit not being exactly in the correct position, and to small movements of the slits and reflecting prisms.

(3) Four of the marks chosen are at the beginning of morse dots or dashes, and four at the end. Averaging the two sets eliminates certain errors inherent in the reception of wireless signals, which are discussed below. To simplify the reduction the four sharp marks are always taken at the start of morse marks and the diffuse ones at the end.

The corresponding points on the base station record are then marked.

Let X, fig. 5, *a*, be one of the chosen points on the field station record, and let A and C be the nearest pendulum marks. Then the distances AX and AC are measured with a measuring microscope and expressed as fractions of a swing by division by ½PQ. The fractions AC* for the 8 points are averaged giving the initial difference in phase between the two field station pendulums.

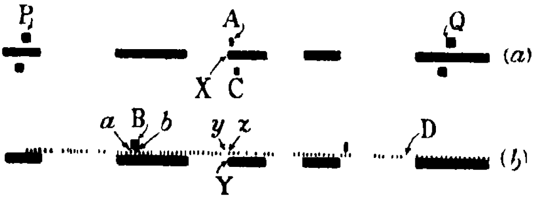


FIG. 5.—Portion of a pair of records. (*a*) Field station, (*b*) base station.

Let Y, fig. 5, *b*, be the point on the message recorded at the base station corresponding to X on the field station record, and B the nearest pendulum mark. Let *yz* and *ab* be the 1/100 second marks between which Y and B lie.† Then *a*B and *ab* are measured and the former expressed as a fraction of the latter, *y*Y and *yz* are similarly treated; also the number of 1/100 second marks between *a* and *y* is counted. The readings are all taken to 0·01 mm., the time represented by BY is now known in terms of the frequency of the supply mains. YB cannot exceed about ¼ second and tests at various times have shown that the error introduced by assuming the frequency to be equal to its nominal value of 50 does not exceed 10^{-4} seconds; thus BY is known to this accuracy in seconds, and, by assuming an approximate value for the period of the base station pendulum, its value in pendulum swings may be found.

Assuming the chosen morse marks to have been made simultaneously on the field and base station records, we may, by subtracting AX from BY obtain the fraction of a swing of the base station pendulum‡ by which the field pendulum is ahead of the base pendulum. This is repeated for each of the 8 chosen points.

The results are arranged in two groups, four points with "sharp" pendulum marks on the field record, and four with "diffuse" marks. The difference in phase between the field and base station pendulums changes steadily owing to the difference in their periods. The approximate rate of change is calculated from the points in one group; and all are reduced, using this rate of change, to the value they would have given if made at the mean time of the eight points. The values in each group of four should then agree, providing a check on gross errors. The average of all eight points is then taken, this average is independent of errors in the assumed rate of change of phase.

---

* These are practically constant for each group of four points, as the two field pendulums only differ in period by about 10^{-6} seconds.

† It is necessary to use the 1/100 second marks to obtain YB as a fraction of a swing, since the motion of the paper is not sufficiently constant for the method employed for the much shorter times AX and AC measured on the field station record to be used.

‡ The times of swing of the pendulums are so nearly the same that it is immaterial that the short interval AX is measured in terms of the field station pendulum.

The pair of records taken at the end of the swing is then treated in a similar manner Subtracting the result obtained from that given by the first pair gives the fractional part of a swing gained or lost by the field pendulum during the hour. The whole number of swings gained is obtained as follows. The number of swings between two marks made by the counter on the base station records, as at D, fig. 5, is known. Thus the number of swings between the two mean points to which the differences of phase have been reduced is known, and, using the approximate rate of change of phase determined above, the whole number can be calculated. This whole number is always small (for the pendulums used about 10) and can be fixed with entire certainty.*

Combining this whole number with the fraction already obtained and dividing by the total number of swings we obtain a more accurate value for the rate of gain of the field station pendulum, and, using the approximate period of the base station pendulum, we can calculate the difference of period between the field and base pendulums. This is reduced, by means of experimentally determined coefficients, to what it would be if both pendulums were swinging with an indefinitely small arc in a vacuum at 0° C.

From the change in the difference of phase between the two field station pendulums during the swing, the difference in period between them is found, and, using the difference between the base and first field station pendulum, the difference between the base and the second field pendulum is found, and finally the difference between the base and the mean field pendulum. Two complete determinations are made at each station; from the mean of these, and the corresponding mean when all the pendulums are swung together at the base station† the change in the mean period of the field pendulums on being taken from the base to the field station is found, and thence the difference in g.

---

* By the difference in appearance of alternate pendulum marks it is at once obvious whether the whole number is odd or even, which provides some check on the arithmetic.

† Three observations are made at the base station before observing at the field station and three afterwards. The mean of all six observations is used.

A slightly abridged example of the calculation for one station is given in Table I.

The calculations can be very greatly shortened by the use of a special measuring microscope giving AX and AC, fig. 4, as fractions of PQ, and *a*B and *y*Y as fractions of *ab* and *yz* directly. They could be further reduced by employing a tuning fork whose frequency is 1/100 of a pendulum swing instead of a vibrator driven by the supply mains.*

6. *Discussion of Errors.*

The errors in a gravity observation are of two kinds —

(*a*) Errors in observing the time of swing of the pendulums.

(*b*) Factors other than changes in gravity causing the time of swing to vary (such as changes in the length of the pendulums, and motions of the ground).

Only the first class are discussed in this section.

The errors in observing the time of swing may be classed as :—

(*a*) Errors in measuring the lengths on the records

(*b*) Errors in converting these into time.

(*c*) Variable lags in the recording system.

These will be discussed separately.

(*a*) By repeating the measurements on a pair of records after an interval of several days the probable error of setting on the centre of a pendulum mark or on the start or end of a morse mark was found to be about 0·01 mm. (0·00014 second), and the probable error in setting on 1/100 second mark considerably less than this. It may be shown that these uncertainties will introduce a probable error of 4×10^{-4} second into the phase difference between the two pendulums as determined from a single point on the records. The uncertainty in the time of swing derived from the mean of the eight points at the start and eight at the end of 1 hour, containing about 7200 swings, is $0 \cdot 3 \times 10^{-7}$ second. The probable error produced in the difference of period of the two field pendulums is $0 \cdot 2 \times 10^{-7}$ second.

(*b*) Irregularities in the running of the field station chronograph drum can produce errors. These are most easily investigated by grouping the residuals

---

* An accuracy of 1/2000 in frequency giving a maximum error of 0·0001 second would be amply sufficient.

Table I —Royston, October 29, 1932 Reduction of Record taken at start of 1st Swing (see fig. 5).

	No. 1	No 2.	No. 3	No 4.	No. 5	No 6.	No. 7.	No 8.	
AX	−0·04	−0·61	−0·05	−0·39	+0·04	+0·30	−0·60	−0·54	mm
AC	−0·43	−0·44	−0·41	−0·36	−0·47	−0·46	−0·40	−0·39	mm
PQ	67·8	67·9	68·1	68·1	67·8	68·0	68·0	69·4	mm
AX	−0·0012	−0·0180	+0·0015	−0·0115	+0·0012	−0·0088	−0·0176	−0·0155	½-swings
AC	−0·0127	−0·0130	−0·0120	−0·0106	−0·0139	−0·0135	−0·0118	−0·0112	Mean = 0·0123 ½ swings
aB *yY*	0·62 0·44	0·48 0·35	0·18 0·36	0·57 0·66	0·42 0·53	0·06 0·38	0·59 0·23	0·46 0·57	mm.
ab *yZ*	0·73 0·67	0·69 0·68	0·70 0·67	0·66 0·70	0·68 0·67	0·66 0·66	0·64 0·68	0·64 0·69	mm
aB *yY*	0·85 0·66	0·70 0·51	0·26 0·54	0·86 0·94	0·62 0·79	0·09 0·58	0·92 0·34	0·72 0·83	0·01 sec.
ay	11	16	16	20	12	12	19	15	0·01 sec.
BY	10·81	15·81	16·28	20·08	12·17	12·49	18·42	15·11	0·01 sec
BY	−0·2139	−0·3128	−0·3220	−0·3973	−0·2408	−0·2472	−0·3644	−0·2990	½-swings*
AX − BY	+0·2127	+0·2948	−0·3235	−0·3858	+0·2420	+0·2560	+0·3468	+0·2835	½-swings*
Number of swing	173	117	99	61	146	138	80	122	Mean = 100
	Approximate rate of change of AX − BY = (0·2127 − 0·3858) (173 − 61) = 0·00156 swings per swing								
Reduced AX − BY	−0·3244	−0·3208	+0·3220	+0·3245	+0·3135	+0·3149	−0·3162	+0·3173	½ swings
Mean		+0·3229				+0·3155			Mean = − 0·3192 ½ swings
Residuals	−15	−21	−9	−16	−20	−6	+7	+18	10^{-4} ½-swings

* Period of base station pendulum = 0·50545 sec.

Calculation of *g*

	Field pendulum A − Base pendulum	Field pendulums A − C.	
Mean phase difference at start (see above)	+0·3192	−0·0123 ½-swings	No of swings = 8900
Mean phase difference at end (from similar calculations)	+0·2720	−0·0331 ½-swings	Therefore approximate No gained
			= 8900 × 0·00156
Fractional part of No of swings gained	−0·0472	+0·0331 ½-swings	= 13·9
Total	14·0472	+0·0331 ½-swings	
Rate of gain	0·00157834	0·00000372 swings/swings	
Difference in period	7977·7 × 10^{-7}	+18·8 × 10^{-7} sec	
Correction for pressure, temperature and arc	−3·8	−0·2	

Reduced difference in period	A 7973·9	A−C 18·6	C 7955·3	Mean	7964·6 × 10^{-7} sec.
Reduced difference in period from second swing					7960·3 × 10^{-7} sec
				Mean	7962·4 × 10^{-7} sec.
Reduced difference in period with both sets at base station					7900·1 × 10^{-7} sec.

Change in period	62·3 × 10^{-7} sec.
Therefore difference in *g*	−0·0242 cm./sec.^{2}
g at base station	981·265 cm./sec.^{2}
Therefore *g* at Royston	981·241 cm./sec.^{2}

of single measurements of the time interval between the passage of the two field station pendulums through their equilibrium positions according to the length of the interval. The residuals of the two groups of four measurements on each record were found, and the probable errors calculated. The mean values of these are shown in fig. 6 plotted against the length of the interval. The probable error is seen to be about 0·0002 second when the length measured is zero, and to increase by about 0·0002 second for each 0·01 second in the interval measured The former figure is just what would be expected from the uncertainty in the microscope settings discussed above, the latter is the result of irregularities in the motion of the recording paper. The points measured on the field records are chosen so as to make the distances measured as small as possible. The average for 18 records was 0·02 second, and the greatest 0·06 second. The probable error in the phase difference of the two pendulums determined from a single point due to this cause will therefore be about 0·0004 second. By a similar argument irregularities in the base station chronograph will produce probable errors of about 0·0002 second. The combined effect on the difference of period of field and base station pendulums determined from the mean of eight points will be $0{\cdot}3 \times 10^{-7}$ second.

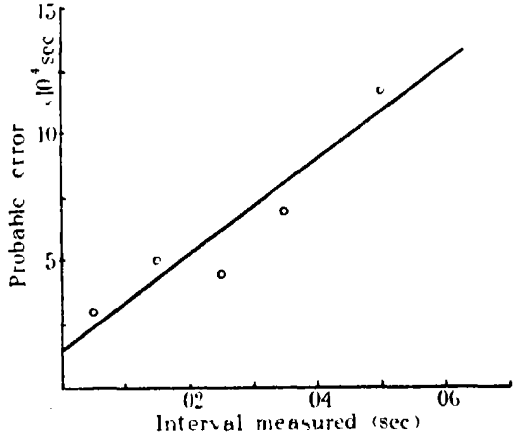


FIG 6 - Correlation of residuals with measured interval

(*c*) *Lags in the Recording System.*—It is well known that if a periodic driving force of the form shown in fig. 7, *a*, be applied to any resonant system the motion produced will not commence or end sharply, but will be of the form shown in fig. 7, *b*. If the equation of motion of the system be

$$\begin{aligned} x + 2kx + \omega^2 x &= 0 \qquad t < 0 \\ &= \mathrm{F} \sin pt \qquad t > 0 \end{aligned} \qquad (1)$$

$\omega^2 x$ being the restoring force, $2kx$ the frictional force, and $F \sin pt$ the impressed force (all per unit mass), then it is easily shown that, if $\frac{\omega - p}{\omega} \ll 1$, the amplitude of the resultant disturbance grows as

$$\left.\begin{array}{ll} & x_0(1 - e^{-kt}) \\ \text{and decays as} & x_0 e^{-kt} \end{array}\right\} \qquad (2)$$

The maximum amplitude x_0 is

$$x_0 = \frac{X}{\sqrt{1 - \left(\frac{\delta f}{k}\right)^2}}, \qquad (3)$$

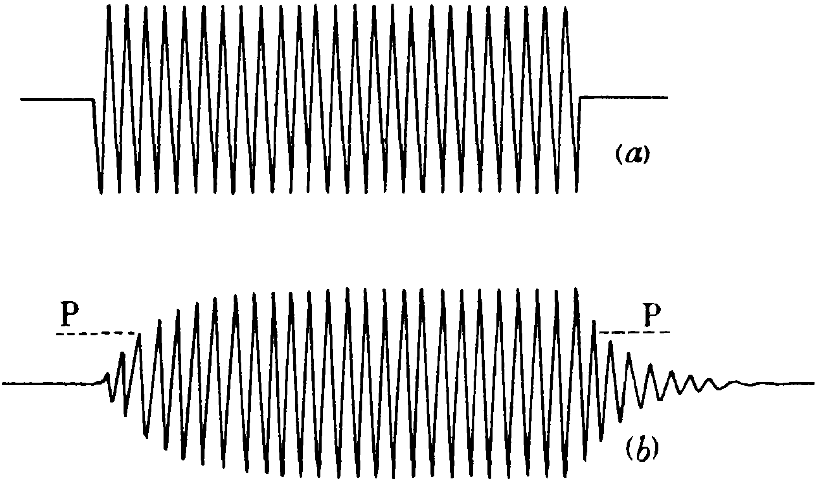


Fig. 7.—Effect of resonant system on signal.

where X is the amplitude when the applied force has a frequency equal to the natural frequency of the system, and $2\pi\delta f$ is the difference between the actual frequency of the applied force and the natural frequency. The results for more complicated systems will be similar.

Now the radio-and audio-frequency tuned circuits of a wireless transmitter and receiver are analogous to the system discussed above, and are governed by equations of the form (1), thus the response to a morse dot will be as shown in fig. 7, *b*. Since a change in the signal cannot be detected till it exceeds the background noise present, a lag will be introduced depending on (*a*) the rate of change of the receiver output due to the wanted signal, and (*b*) the amount of interference from other stations. From (2) the maximum rate of change of the receiver output is kX, the receiver being tuned to the signal From (3) response to an interfering signal, which would give a response X' at resonance, is $X' \Big/ \sqrt{1 + \left(\frac{\delta f}{k}\right)^2}$. If, as is clearly desirable, the selectivity is sufficient to reduce the interfering signal to a small fraction of its maximum value, this becomes $kX'/\delta f$. Thus the time before the signal becomes appreciable, will be $kX'/\delta f \div kX = X'/X\delta f$, which depends on the signals only, and not at all on the receiver. The argument is quite general, and independent of the particular receiving arrangements employed; the higher the selectivity the slower is the rise and fall of the signal, on the other hand the less the selectivity the more the interference and the more the signal has to grow before becoming appreciable.*

The errors produced may be reduced :—

(*a*) By taking the initial and final records under exactly the same conditions, thus keeping the lag constant, for it is only a time interval and not absolute time that is required.†

(*b*) By arranging the reduction of the records so as to eliminate changes of lag due to small changes of adjustment. Let P, fig. 7, *b*, represent the position of the chronograph slit, then if the zero line moves up a distance ε relative to this slit the time of first passage of the light across it will become earlier, and the time of last passage at the end of the dot later. The change will be $\varepsilon \Big/ \frac{dx}{dt}$ in each case, or, from (2) above

$$\frac{\quad}{k\,(x_0 - x)} \qquad \text{and} \qquad -\frac{\varepsilon}{kx},$$

these are equal and opposite if $x = \frac{1}{2}x_0$, that is, if the slit is placed half way between the position of the light with no signal and its extreme position with a signal. A similar compensation exists for changes of signal strength. The only change which is not compensated is a change in the decrement of the circuits, and as no reaction control is provided, this seems most unlikely to occur. Thus, if half the measurements be taken on the beginning of morse signals and half on the end, errors due to small changes in adjustment of the oscillograph and receiver will be eliminated. As it is convenient to have the interfering signals reduced to much less than half the wanted signal, this method of working will lead to a total lag several times the theoretical $X'/X\delta f$. As the frequency spacing of stations in this wave band is about 200 cycles, this theoretical minimum is 0·001 second for equal "in tune" strengths of the wanted and unwanted signals.*

---

* See 'Radio Research Board Special Report No. 12,' "High Selectivity Tone Corrected Receiving Circuits" for a discussion of the matter from a rather different point of view.

† It seems possible that the different values given by different observatories for the interval between the same pairs of time signals may be due to neglect of this precaution. For example, if the interval between a strong signal and a weaker one be determined in terms of an observatory clock, and the strength of the weaker signal be increased by the use of reaction, an error of the actually occurring order of magnitude will be introduced, owing to the change in damping of the circuits (k in equation (1)).

In order to check these conclusions oscillograms of the motion of the spot of light were taken, these were found to be of the required form, and in particular the rate of rise and fall of the signal was closely the same at half the maximum signal. From such records the natural periods of the oscillographs were found to be 1/900 second, the decrement k to be 80 sec.$^{-1}$, and the time to rise to half the maximum signal 0·008 second.

In order to see how large were the changes which had to be eliminated by the above procedure, the lengths of a large number of dots were measured on both the field and base station records, and expressed in seconds. The mean change in the difference between these on the field and base station records which occurred between the initial and final records of a swing gives twice the error to be eliminated. The value obtained for the error was 0·002 second, it therefore seems probable that the actual uncertainty produced with the procedure used will not be more than 1/10 of this, producing an error of not more than $0\cdot3 \times 10^{-7}$ second in the period or 1×10^{-4} cm./sec.^{2} in g.

The oscillatory motion of the recording spot of light produces small random errors. For if at the end of one oscillation it just fails to reach the slit, no mark will be made till one cycle later. The error is easily seen to be random, and to produce a probable error of half the period of vibration. The oscillographs used have periods of 1/900 second, so that the probable error produced is 0·00055 second for each oscillograph, or 0·0008 second for the total effect on the phase measured from a single point. The effect on the period determined from the mean of eight points at the start and end of the swing will be $0\cdot5 \times 10^{-7}$ second.

A correlation of the residuals of the individual measured points with the fractions of a swing interpolated on the base station record was looked for but not found, showing that the irregularities in the 1/100 second marker are negligibly small.

The errors discussed are summarized in Table II. In addition there is the error of not more than $0{\cdot}3 \times 10^{-7}$ second produced by the lags in the wireless receiver, which is systematic for the points of a single swing, but random for different swings.

---

* The lag could, of course, be reduced by working with less sharply tuned circuits on a less crowded wave band, but the Rugby signals are very convenient in practice.

Table II.—Probable Errors.

Cause.	Field-Base.		Difference in period of 2 field pendulums
	Phase difference from single point	Final difference in period	
	10^{-4} sec.	10^{-7} sec.	10^{-7} sec.
Measurement of records	4	0·3	0·1
Irregularities of motor	4	0·3	0·3
Recording of wireless signals	8	0·5	
Combined effect	10	0·7	0·3

The probable error of the difference of phase between the field and base pendulums determined from a single point on a pair of records would therefore be expected to be about 0·0010 second. Comparison of this with the value obtained from the mutual consistency of the individual points on a record provides a very stringent check on the completeness of the above investigation of the random errors. By forming the residuals of each group of four measurements of the phase difference, computing the probable error by Peter's formula, and averaging for all the groups available, the value 0·0011 second was obtained. The frequency distribution of the 512 residuals used in obtaining this probable error is shown in fig 8, the curve being that required by theory for this probable error, the agreement with the normal curve is probably as good as could be expected with the limited number of observations available.

Thus it may be concluded that the probable error in the determination of the difference in period between the field and base station pendulums is about $0{\cdot}7 \times 10^{-7}$ second (corresponding to $0{\cdot}3 \times 10^{-3}$ cm./sec.^{2} in g) for a 1-hour swing or $0{\cdot}5 \times 10^{-7}$ second ($0{\cdot}2 \times 10^{-3}$ cm./sec.^{2}) for the mean of two swings; which is less than the variation in the period from causes other than changes in gravity. Further, it is clear from the discussion given above, that there would be no great difficulty in reducing the error should this be desirable.

7. *The Observations.*

In order to test the practicability of the method for field measurements and to reveal unsuspected sources of error, two sets of observations were made connecting three stations near Cambridge with the "Pendulum House" in the grounds of the University Observatory. Each consisted of two 1-hour swings, and each had its own base station measurements; an interval of a fortnight was left between the two sets. The differences between them should fairly represent the capabilities of the method. The results are given in Table III. The probable error of each set calculated from the three differences, is 1×10^{-7} second in the mean period, or 0·0004 cm./sec.^{2} in g. A closer examination of the agreements of the individual swings shows that this estimate is probably fortuitously low, the true value being more probably about 0·001 cm./sec.^{2} In addition, a set of measurements was made connecting the Pendulum House with the Cavendish Laboratory, where g was found to be 0·0018 cm./sec.^{2} more than at the Pendulum House. Table V gives the

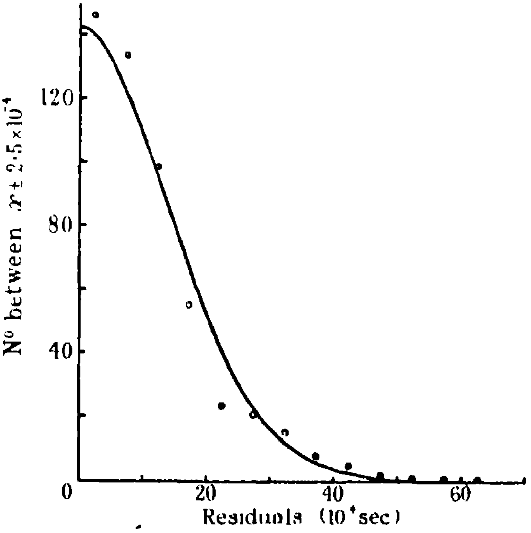


Fig. 8.—Frequency distribution of residuals.

Table III.

	$g - g$ (base station) 10^{-3} cm./sec.^{2}		
	Waterbeach	Granchester	Royston
1st set	+8·4	−1·8	−24·2
2nd set	+7·5	−2·4	−23·7
Mean	+8·0	−2·1	−23·9

g at the base station is 981·265

results of the individual swings in the three sets of measurements. The probable errors computed from the differences between the results of individual swings taken on a single visit to a station are given in Table IV.

Table IV.

Pendulum "A."	Pendulum "C."	Mean $\frac{1}{2}$(A + C)	Difference A − C.
seconds $3{\cdot}1 \times 10^{-7}$	seconds $3{\cdot}0 \times 10^{-7}$	seconds $2{\cdot}9 \times 10^{-7}$	seconds $2{\cdot}0 \times 10^{-7}$

This probable error will produce a probable error of 0·0012 cm./sec.^{2} in the value of g deduced from a single swing, or 0·0008 cm./sec.^{2} on the value deduced from the mean of two 1-hour swings, in addition we must allow $0{\cdot}0012/\sqrt{6}$ for the uncertainty in the mean of the six base station measurements, giving 0·001 cm./sec.^{2} as the probable error of the final result, as stated above. This is about twice the value deduced from the consistency of the final results of the two visits to three of the stations. The uniformity of the "Bouger" anomalies given in Table VI gives added confidence in the accuracy of the results.

The observed errors given in Table IV are all considerably larger than the estimated uncertainties in observing the periods; moreover, the probable error of "A–C" is nearly as large as that in the periods of the pendulums themselves, although it is immune from nearly all the possible sources of observational error.* It may, therefore, be concluded that the major part of the error is due to real variations in the time of swing of the pendulums. As most of the sets of observations from which these probable errors are deduced were completed in a few hours, there seems little likelihood that changes in the lengths of the pendulums can cause any large part of the observed irregularity. The most probable cause seems to be motions of the ground. If the motion of the ground at the base station produces an error ε_1, and that at the field station ε_2, then the resultant errors are :—

Pendulums "A"—"Base."	Pendulums "C"—"Base."	Mean $\frac{1}{2}$("A"+"C") —"Base."	Difference "A"—"C."
$\sqrt{\varepsilon_1^2 + \varepsilon^2_2}$	$\sqrt{\varepsilon_1^2 + \varepsilon_2^2}$	ε_1	$2\varepsilon_2$

---

* In particular from all errors in recording the wireless signals.

Since the mean of the two field station pendulums is independent of the motions of the ground at the field station,* and the difference of the two is independent of motions occurring at the base station. If $\varepsilon_1 = 3 \times 10^{-7}$ and $\varepsilon_2 = 1 \times 10^{-7}$ second we get the four observed probable errors, showing that the explanation is a possible one. If a pendulum is left hanging at rest on its knife edges at the base station, it is found after a few minutes to acquire an amplitude of a few minutes of arc, showing that the motion of the ground is of the order required to produce the observed errors. In future two pendulums will be used at the base station to eliminate this source of error.

Examination of the observed mean difference in period on the various visits to the base station shows changes which are too large to be ascribed to observational error, and are presumably due to changes in the pendulums.† For each of the three field stations and for the base station the difference of period between each field pendulum and the base pendulum is greater for Set 2, Table V, than for Set 1.

8. *Effect of Sway of the Support.*

If the two field station pendulums were exactly similar, and were started with exactly the same amplitude, and in exactly opposite phase, no uncertainty would be produced by yielding of the pendulum support under the forces produced by the swinging pendulums.

In practice these conditions are not exactly fulfilled. The two field pendulums differ in period by about 15×10^{-7} second, and from the mean of a number of swings the mean difference in initial amplitude is 4 per cent., and the mean initial difference in phase 4° (corresponding to 1′ in the actual position of the pendulums with the normal amplitude of 0·4°). The effect of these on the periods of the individual pendulums, when the "sway correction" for one pendulum swinging alone is 40×10^{-7} second, may be shown‡ to be 2×10^{-7} second, and on the mean pendulum $< 10^{-8}$ second, which is negligible. The yielding of the base station pendulum support is very small and constant.

---

* Vening Meinesz, "Theory and Practice of Pendulum Observations at Sea," Delft, 1929.

† Unfortunately none of the pendulums were observed against time signals before and after these observations, so that it is not possible to say which one has changed. It seems desirable to check the base pendulum (which always hangs on its knife edges) about once a week.

‡ Andersen, 'Beitr. Mitschwingungstheorie,' København, 1932, gives a convenient discussion of all the relevant formulæ.

Table V.

Station.	Date	Difference of period, Field − Base.				
		Pendulum A	Pendulum C	Mean ½(A + C).	Residuals For ½(A+C)	A−C.
		Set 1.				
		10^{-7} sec	10^{-7} sec	10^{-7} sec	10^{-7} sec	10^{-7} sec.
Pendulum House	14 10 32	7909·5	7893·3	7901·4	+5·0	+16·2
	,,	7094·4	7884·0	7894·2	−2·2	+20·4
	15 10 32	7903·3	7883·7	7893·5	−2·9	+19·6
	Mean	7905·7	7887·0	7896·4		+18·7
Waterbeach	18 10 32	7885·0	7874·0	7879·5	+1·1	+11·0
	,,	7884·8	7869·6	7877·2	−1·2	+15·2
	Mean	7884·9	7871·8	7878·4		+13·1
Grantchester	19.10 32	7919·5	7901·0	7910·2	+5·4	+18·5
	,,	7910·7	7888·4	7899·5	−5·3	+22·3
	Mean	7915·1	7894·7	7904·8		+20·4
Royston	29 10 32	7973·9	7955·3	7964·6	+2·2	+18·6
	,,	7968·3	7952·3	7960·3	−2·1	+16·0
	Mean	7971·1	7953·8	7962·4		+17·3
Pendulum House	1 11 32	7908·1	7894·5	7901·3	−2·6	+13·6
	,,	7911·5	7895·1	7903·3	−0·6	+16·4
	2 11 32	7914·0	7900·3	7907·2	+3·3	+13·7
	Mean	7911·2	7896·6	7903·9		+14·6
		Set 2.				
Pendulum House	14 11 32	7921·0	7907·1	7914·0	+1·2	+13·9
	15 11 32	7915·2	7908·1	7911·6	−1·2	+ 7·1
	Mean	7918·1	7907·6	7912·8		+10·5
Waterbeach	16 11 32	7887·5	7886·2	7886·9	−2·9	+ 1·3
	,,	7897·5	7887·8	7892·7	+2·9	+ 9·7
	Mean	7892·5	7887·0	7889·8		+ 5·5
Grantchester	17.11 32	7909·9	7909·9	7909·9	−5·4	0·0
	,,	7923·4	7917·7	7920·6	+5·3	+ 5·7
	Mean	7916·6	7913·8	7915·3		+ 2·8
Royston	18 11 32	7978·1	7974·0	7976·1	+5·8	+ 4·1
	,,	7966·1	7962·9	7964·5	−5·8	+ 3·2
	Mean	7972·1	7968·4	7970·3		+ 3·6
Pendulum House	22 11 32	7907·2	7906·2	7906·7	+1·1	+ 1·0
	23 11 32	7903·2	7903·0	7903·1	−2·5	+ 0·2
	,,	7911·0	7903·2	7907·1	+1·5	+ 7·8
	Mean	7907·1	7904·1	7905·6		+ 3·0

Table V—(continued).

Station.	Date.	Difference of period, Field — Base.				
		Pendulum A	Pendulum C.	Mean ½ (A+C).	Residuals for ½ (A+C).	A−C.
		Set 3.				
		10^{-7} sec	10^{-7} sec	10^{-7} sec.	10^{-7} sec	10^{-7} sec.
Pendulum House	15 9 32	7900·1	7885·9	7893·0	−2·3	+14·2
	17 9 32	7904·6	7890·7	7897·6	+2·3	+13·9
	Mean	7902·4	7888·3	7895·3		+14·0
Cavendish Laboratory	26 9 32	7900·4	7883·9	7892·2	−0·5	+16·5
	,,	7905·0	7891·7	7898·3	+5·6	+13·3
	,,	7897·7	7884·9	7891·3	−1·4	+12·8
	27 9 32	7896·4	7881·5	7889·0	−3·7	+14·9
	Mean	7899·9	7885·5	7892·7		+14·4
Pendulum House	30 9 32	7907·9	7894·6	7901·2	+2·5	+13·3
	,,	7900·8	7891·2	7896·0	−2·5	+9·6
	Mean	7904·3	7892·9	7898·6		+11·4

N.B.—Pendulum "C" was accidentally reversed throughout Set 2. This does not in any way effect the results but renders the values of "A — C" in this set not comparable with those in the other two sets.

9. *Corrections for Variation of Temperature and Pressure.*

The variation of period with temperature and pressure was determined by setting up one pair of pendulums in a room, which could be heated, and recording their swings on the photographic chronograph described above. The swings of a reference pendulum, kept at constant temperature and pressure in another room, were also recorded by means of a photoelectric cell coupled, through an amplifier, to an oscillograph. The temperature coefficients of the times of swing are all in the neighbourhood of 3×10^{-7} sec./° C.*.

10. *Further Developments.*

It is hoped in the near future to develop the method in two directions.

(*a*) To shorten the time spent at each station. It is at present necessary to leave the pendulums in the case in which they are to be swung for about

---

* The value $0 \cdot 5 \times 10^{-7}$ sec./° C. given by Lenox-Conyngham (*loc. cit.*) p. 328, is erroneous.

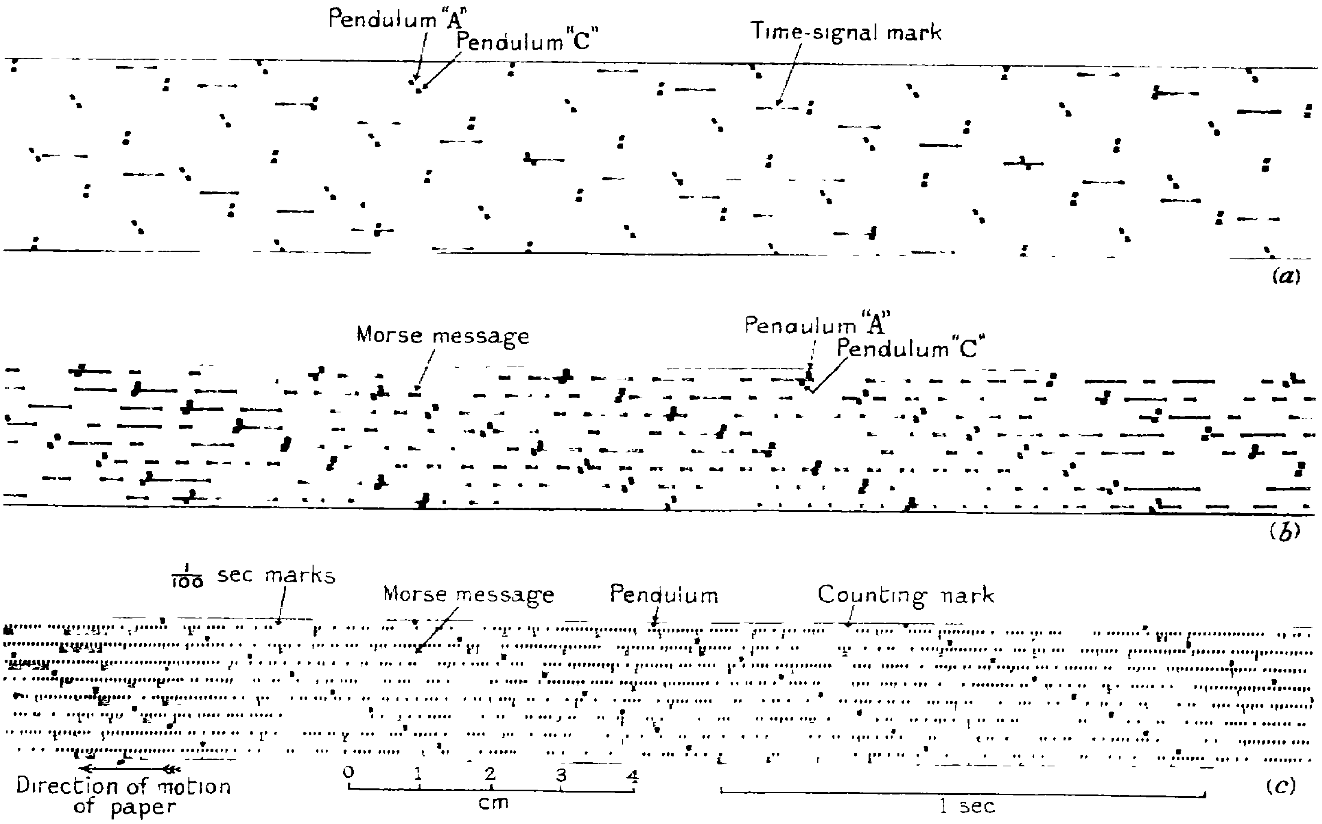


FIG. 9.—Small portions of records. (*a*) Time signal and two pendulums. (*b*) Morse message and two pendulums (field station). (*c*) Morse message and 1/100 sec marks (base station).

Table VI.

Station.	Latitude ϕ	Longitude λ	Height H	Observed g_0	Theoretical free air value γ_A *	Free air anomaly $g_0 - \gamma_A$	Bouguer anomaly †
	° ′ ″	° ′ ″	metres				
Pendulum House	52 12 52	0 05 44 E	25	(981 265)	981 254	+11	+9
Waterbeach	52 15 30	0 10 50 E	6	981 273	981 264	+ 9	+8
Grantchester	52 11 10	0 05 10 E	19	981 263	981 254	+ 9	+7
Royston .	52 02 50	0 02 10 W	67	981 241	981 227	+14	+8
Cavendish Laboratory	52 12 10	0 07 20 E	11	981 267	981 258	+ 9	+8

* $\gamma_A = 978\ 030\ [1 + 0{\cdot}005302 \sin^2 \phi - 0{\cdot}000007 \sin^2 2\phi] - 0{\cdot}3086 \times 10^{-3}\ \mathrm{H}$ (Helmert, 1901)

† Bouguer anomaly = Free air anomaly − 0·000086 H Density of sub-soil 2·1

2 hours to allow them to acquire the same temperature as the thermometer.* This could be avoided by carrying the pendulums in the case in which they are swung.

(*b*) To decrease the uncertainty produced by changes in the length of the pendulums when carrying them from place to place. It is hoped that the use of quartz pendulums of the Kohlschütter "Minimum" form† will be an improvement.

In conclusion I wish to thank Sir Gerald Lenox-Conyngham, F.R.S., under whose direction this work was carried out, for his advice and encouragement.

11. *Summary.*

A method of comparing the periods of pendulums swinging simultaneously at two stations is described, This does not require special wireless signals, and enables a relative determination of g to be made with an accuracy of 1/1,000,000 in two swings of 1 hour each. A method of comparing gravity pendulums and wireless time signals is also described.

---

* In this time any initial temperature difference has sunk to one-third of its initial value if the case is evacuated, or to one-seventh if it is not

† Kohlschutter. 'Z Geophys,' vol 6, p 466 (1930)

Experiments on the Transmutation of Elements by Protons.

By M. L. E. OLIPHANT, Ph.D., Messel Research Fellow of the Royal Society, and Lord RUTHERFORD, O.M., F.R.S., the Cavendish Laboratory, Cambridge.

(Received June 16, 1933.)

Introduction.

In the pioneer experiments of Cockcroft and Walton* on the transformation of elements by bombardment with swift protons, it was found that easily measurable disintegration effects could be observed in certain elements for comparatively small accelerating voltages—of the order of 100,000 volts. Their experimental arrangement was specially devised for the use of much higher voltages. The proton stream from the discharge tube passed through a long evacuated path before reaching the target, and spread over a considerable area. Proton currents of the order of 1 microampere were usually employed. In order to study these transformation effects at still lower voltages we constructed a special discharge tube which would give us a more intense and concentrated stream of protons, of the order of 100 microamperes, using an accelerating voltage not greater than 250,000 volts. The stream of charged particles generated by the discharge through hydrogen was analysed into its components by means of a magnetic field, so that the effects due to the particles H^+ and H_2^+ could be separately determined. The stream of protons fell on a target of about 1 sq. cm. in area and the solid angle over which particles passed into the detecting chamber was much increased. In these ways it was found possible to obtain for examination at least 1000 times the number of disintegration particles at any voltage as in the original apparatus of Cockcroft and Walton.

This large increase in the number of particles available for counting made it possible to examine the disintegration of such active elements as lithium and boron at low voltages and in the form of very thin films. In addition, it was not difficult to obtain intense but narrow streams of particles in order to analyse with more detail the disintegration as regards both range and velocity of the expelled particles.

The general method has proved very convenient for studying the transmutation of elements by protons of energy between 20,000 and 200,000 volts.

---

* 'Proc. Roy. Soc.,' A, vol. 137, p. 229 (1932).

Naturally a number of preliminary investigations were necessary in order to design the best form of discharge tube to give the intense and constant current of protons, so that measurements could be made under satisfactory conditions. The success of such experiments largely depended on the adequacy of the pumping system to obtain steady conditions in the discharge tube. Fast pumping is required and it is desirable to use oil rather than mercury diffusion pumps as mercury seems to promote autoelectronic discharges from the electrodes. We have found that the pumps constructed for us by Metropolitan-Vickers were very suitable for our purpose.

The Apparatus.

Source of Protons.—Various methods have been tried for obtaining a strong beam of protons, but the simplest and most consistent source so far used has been a modified form of the usual canal-ray discharge tube It is not possible to produce a beam of protons which carries more than a small fraction of the current through the discharge itself, so that in order to obtain a large proton current, a heavy discharge current must be used Experiment shows that the maximum current is produced from a discharge running at about 20,000 volts, and hence with discharge currents of the order of 10-100 milliamperes it becomes necessary to dissipate large quantities of heat at the electrodes and walls of the tube. It will be seen later that the source of the protons must be placed inside the vacuum in the accelerating system, and hence there is no cooling of the walls by convection. A porcelain tube used as an ordinary Wien tube, and cooled all over by immersion in a bath of paraffin oil which was circulated by a pump, gave a good beam, but the walls of the tube rapidly became coated with a conducting layer sputtered from the electrodes. The source in use at present overcomes this difficulty without losing the advantages of the Wien form of canal-ray tube.

The perforated cathode, C, fig. 1, is a block of steel bored out as shown in the figure and screwed tightly into the end of a steel tube B, about 1 metre long and 12 cm. internal diameter. At its upper end this tube is fastened into a steel plate which closes the top of the accelerating cylinder, N. The anode is a second steel tube A, supported axially inside the cathode and separated from it by about 4 mm. all round. This tube extends to within a few millimetres of C and at the upper end it is supported by a pair of insulating glass plates as shown. The plug D, which extends through the plates, is bored from a solid piece of steel and is cooled by a rapid stream of transformer oil supplied by a gear-pump through insulating tubes. Hydrogen can be fed in through a small tube K after passing through an adjustable needle valve. All joints at the top between the glass plates and the steel were made vacuum-tight with the special luting compound, Apiezon Q. This tube will deal with large inputs of power, for the energy of the cathode rays is dissipated on the anticathode D, while the cathode C and the steel tubes can run red-hot, if necessary, and dissipate their energy by radiation. The space between the two steel tubes is too small for a discharge to build up there and it concentrates naturally on the hole in the cathode. Power for the discharge is obtained from a 5-kw. alternator driven through a long insulating belt and mounted on an insulating pedestal, together with the necessary transformer and controls and the resistance and condenser for stabilizing the discharge. The rectifier forms part of the general rectifier system as shown in the diagram. The beam from this tube, when fully outgassed by continued operation, is a very nearly parallel streak of luminosity, and the current collected by the Faraday cylinder F into which it can penetrate is of the order of 1 milliampere with 20 milliamperes in the discharge. At first this beam consists very largely of molecular ions,

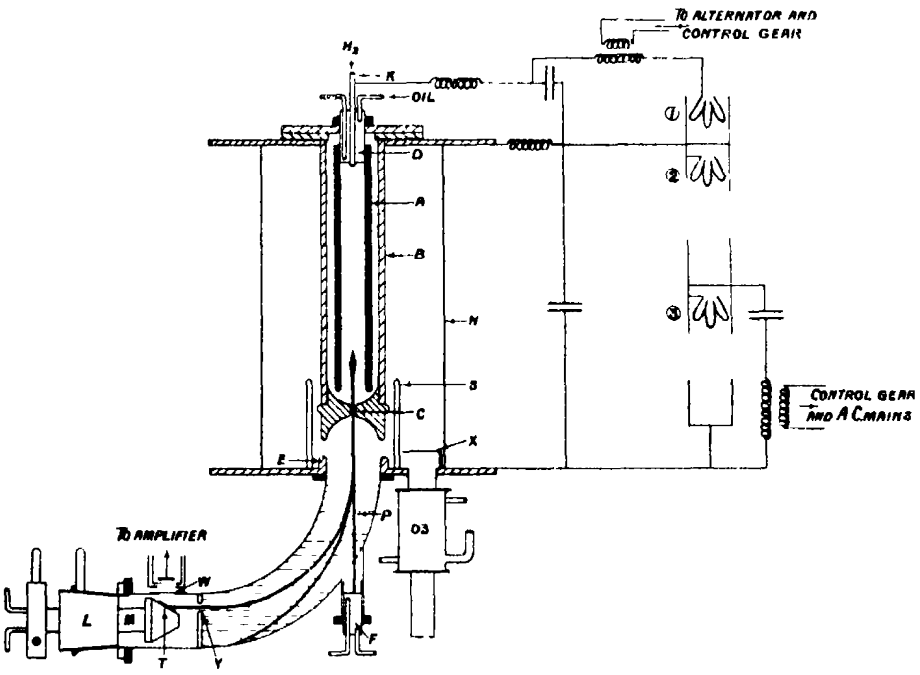


Fig. 1.

but after running for some time it changes over and becomes nearly all protons and neutral particles.

Accelerating System.—The discharge tube is mounted in a tall glass cylinder N, closed at the bottom by a second steel disc. The proton beam is accelerated by a potential of up to 200,000 volts applied between C and E. The steel shield S prevents the impact of secondary electrons upon the glass walls and the consequent heating, while it also serves to remove the strong electric field from the region of the bottom plasticene joint between the glass cylinder and the steel plate. Immediately after acceleration the beam is bent round by a magnetic field produced between the shaped pole pieces P. The mean radius of curvature of the path is 10 cm, and the pole pieces are about 4 cm wide and are separated by a 1-cm gap. The electromagnet used will produce a field of 18,000 gauss between these pole-pieces. A slit Y, 0·5 cm. by 1 cm., serves to define the deflected beam and sort out a given kind of ion with a small range of energies. The beam which passes Y falls upon the target T. This is of steel and has three faces at 45° to the axis, which form a sort of truncated pyramid with a circular base. Each face bears a recess, 1·2 cm. in diameter and about 1 mm. deep. The metal or powder to be bombarded is pressed or hammered into these spaces and the beam strikes the surface which is uppermost. The target is carried on a water-cooled stem M which rotates in a ground joint and can be set so as to bring any one of the three faces into the beam. By having only three faces any one face is completely shielded from the material sputtered from that which is in the beam, and at the same time it is possible to make very rapid comparisons between various targets without disturbing any adjustment, mechanical or electrical. A number of these targets have been made and these can be inserted with great rapidity by unscrewing the joint between M and T. The target system is insulated from the main body of the apparatus by the ebonite sleeve L, and the current to it can be measured with a microammeter. The stray magnetic field over T effectively prevents the escape of secondary electrons, so that the current measured is the true ion current.

Immediately above the bombarded face of the target, and about 1 cm. from its centre, is a mica window W, 1 cm. in diameter and supported on a grid which subtends the smallest possible area to particles coming from the target. The solid angle obtained in this way is approximately 0·7. The windows used have thicknesses varying from 5·0 mm. to 1·2 cm., according to the range of particles under investigation. All particles from the target which penetrate the window can be collected in the chamber of an ionization counter which is mounted immediately above W and which is connected with an amplifier designed for us by Wynn-Williams.

The accelerating chamber is evacuated by an 03 oil condensation pump supplied by Metropolitan-Vickers, and this is screwed directly to the bottom plate of the chamber so that the full pumping speed can be used. The 03 pump is connected directly through wide tubing with an 02 pump, and this is backed by a rotary oil pump of high speed. The shield X is an umbrella-like piece of nickel which prevents a discharge from striking directly down into the diffusion pump if the tube should go soft. The portions of the apparatus which are below E are enclosed in a lead-lined observation space which also acts as an electrostatic shield.

The accelerating voltage is supplied by a transformer giving 100,000 volts. This is rectified and doubled, using the simple circuit developed by Cockcroft and Walton. The rectifiers (2) and (3) are in this circuit and are arranged to give a large current at relatively low voltage drop. The rectifier (1) is used to supply the direct current for running the discharge and is cooled by oil-circulation All three rectifiers are mounted in the same tower of glass cylinders and are evacuated by the same oil diffusion pump. The filaments of rectifiers (1) and (2) are supplied with current through suitable transformers from the alternator which supplies the discharge potential, while (3) is lit by current from a storage battery

A zinc-sulphide counting screen and microscope provided with a right angled prism enable scintillations produced by the particles from T to be observed at any time, so that it is possible to check the fact that the electrical counter is counting genuine particles. Radiation and electrical disturbances are very plentiful in this type of work, and it is essential that this simple check be made periodically, and especially when any doubt arises.

Voltage Measurement.—The total accelerating voltage is measured by a spark-gap as a primary standard, but this is only needed occasionally to calibrate a simple valve-voltmeter, fig. 2 The potential is applied to a plate A placed about 3 feet from an insulated segmented conductor S. In front of this rotates a similar earthed segmented rotor R, so that the vanes on S are alternately exposed to and shielded from the field due to A. R is driven through a spring belt from a synchronous motor, so that the speed is sensibly uniform. S is connected with the grid of one of a pair of valves of relatively low impedance which form two arms of a bridge, as shown in the diagram. The condenser C which is placed in this circuit blocks any charge which collects on S due to corona from A. The valve is used as a leaky grid rectifier and hence gives an almost linear response with suitable values of the automatic grid bias and grid leak. The bridge is enclosed in a suitably screened box and the heating currents for the cathodes and the plate potential are both obtained from the mains in the usual way. A milliammeter connected between the two plates serves to read the voltage, and a series of shunts enables the sensitivity to be changed readily The current passing in the meter is arranged to be small compared with the plate current of either valve, so that an approximately linear characteristic is preserved. The voltmeter is easily calibrated by comparison with the spark-gap at two or three voltages for each range. The sensitivity can be further reduced by sliding the plate P in front of the rotating sectors and shielding off the field of A. In practice S is only about one-quarter exposed A potentiometer altering the potential of the grids enables the zero to be set at any moment with the sectors S earthed. This voltmeter has proved exceedingly reliable and robust, and the zero drift is very small once the valves have reached a steady temperature.

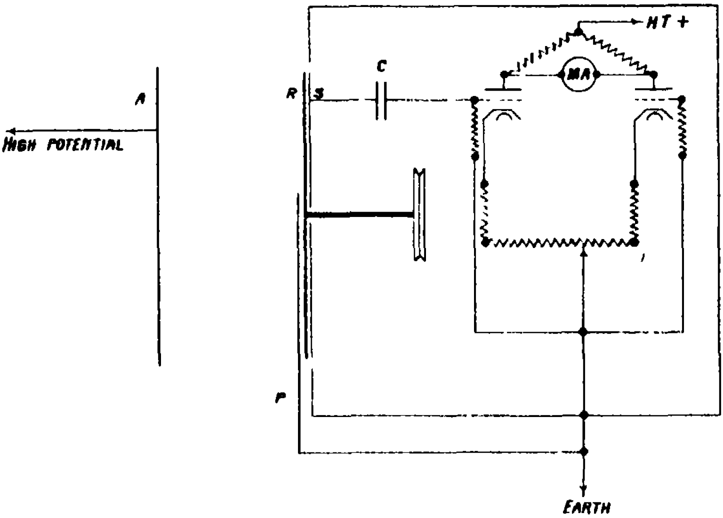


FIG. 2.

Measurement of the Range of the Particles.—The ranges of the particles coming from the target have been measured by observing the reduction in the number counted, as absorbing screens of mica or of gold were placed between the window and the counter. By substituting a weak polonium source for the target a check on the thickness of the absorbing screens and of the "dead-space" between the window and the counter was obtained. We have used also a small absorption chamber fixed above the window and provided with a second window in line with the first. The pressure in this chamber, which was 1·0 cm. deep, could be varied, and the stopping power of the gas calculated. When the number of particles is large this offers the simplest and most direct method for interpolating between mica screens. In some cases more exact measurements of ranges have been made by canalizing the beam of particles from the target by placing stops with small openings over window and counting chamber. This can only be done where the number of particles is large

Experimental Procedure.—The discharge tube was freed from occluded gases by running for a long time, first without and then, as it hardened, with hydrogen flowing in. When it was once degassed it was only necessary to run for about half an hour in the morning to get the tube into a steady state. It was found to be most easy to degas the main accelerating electrodes and tube by admitting hydrogen till it was possible to run a discharge at about 20-60 kilovolts. This is continued at as high a current density as possible for about half an hour and on pumping out the hydrogen it is usually found that the tube is quite hard and stable up to 200,000 volts. The magnetic field is then set for the kind and energy of ion required and the beam is brought on to the target by varying the voltage applied to the main transformer. The particles entering the counter are recorded on an oscillograph or are counted directly by means of the thyratron "scale of two" counter devised by Wynn-Williams.*

The Results.

Lithium.—We have not done much with this element,† but we have confirmed the order of magnitude of the range and the yield found by Cockcroft and Walton In addition we have measured the voltage variation of the number of particles emitted from a thin film of Li in the form of oxide. This was prepared by burning Li in air and holding a piece of cold steel in the ascending stream of vapour for a moment. The film was quite invisible and it is estimated that its stopping power must have been much less than 1 mm. of air equivalent at any point, so that when bombarded it is unlikely that any proton made more than one collision with a lithium nucleus. The voltage variation of the emission of disintegration particles therefore represents the true variation in the probability of disintegration at different voltages. Fig. 3 may thus be said to represent the "voltage disintegration function" for the bombardment of lithium by protons. It is seen that particles begin to be detected at about

---

* 'Proc. Roy. Soc.,' A, vol. 136, p. 312 (1932).

† Since this paper was written lithium has been examined in more detail and the results will be incorporated in a further paper.

40 kv., and that thereafter the numbers increase rapidly but not as rapidly as an exponential. The course of the logarithmic curve indicates that the number of particles may approach a steady value as the energy of the bombarding protons increases above 200 kv., or may even decrease after passing through a maximum.

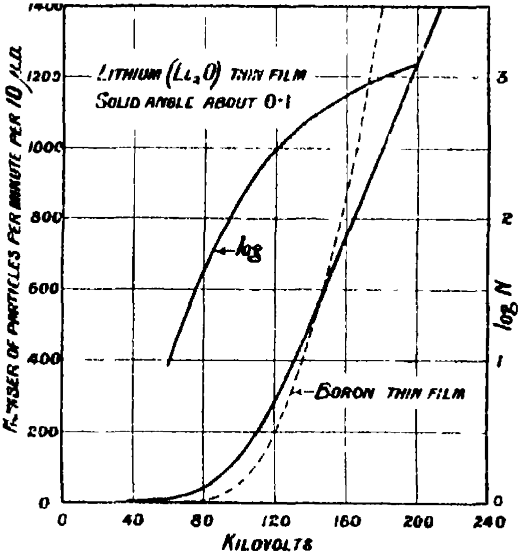


Fig 3

A thick target of Li metal, bombarded by the direct mixed beam, gave a practically exponential rise in the number of emitted particles as the energy of the beam was increased, and particles were detected at voltages as low as 30 kv , fig 4 The currents in some of these experiments were of the order of 50 microamperes and the solid angle over which particles were collected was about 0·7, so that it was possible to detect particles when their total number was only about 1/10,000 of the number present in the experiments of Cockcroft and Walton. Notwithstanding this large factor we have been unable to detect particles with any certainty below about 30 kv.

Boron.—We have determined the form of the voltage variation of the number of disintegration particles from boron for protons bombarding both thick films and thin films of the element. Fig. 5 shows the curve obtained from a thin film of boron prepared by evaporating a single drop of a dilute solution of borax as uniformly as possible over the surface of an iron target, and then heating to a red heat in an atmosphere of coal gas. The weight of borax may be calculated from the strength of the solution, and assuming the

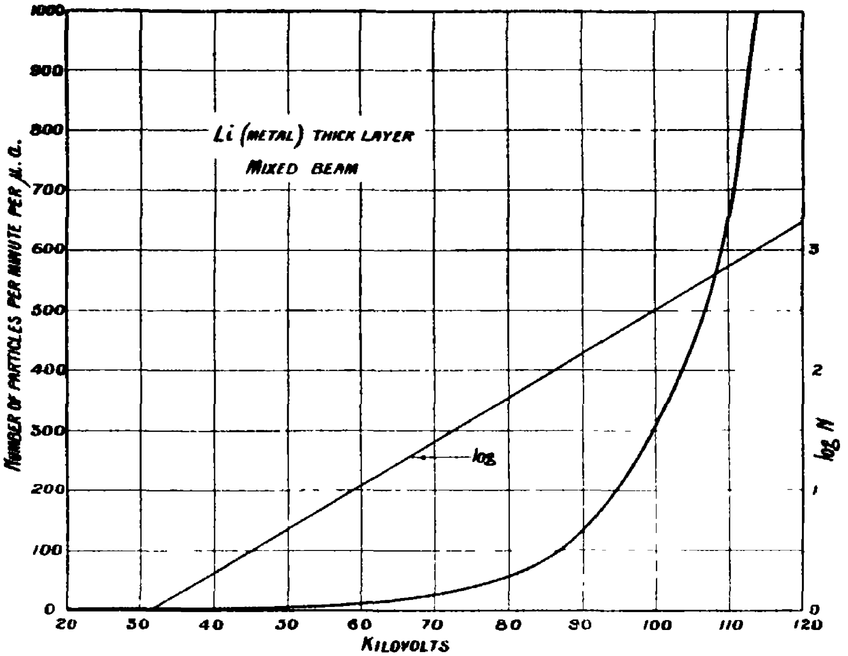

Li (metal) thick layer
Mixed beam
log
Number of particles per minute per μ.a.
Kilovolts
log N
0
100
200
300
400
500
600
700
800
900
1000
20
30
40
50
60
70
80
90
100
110
120
0
1
2
3

Fig. 4.

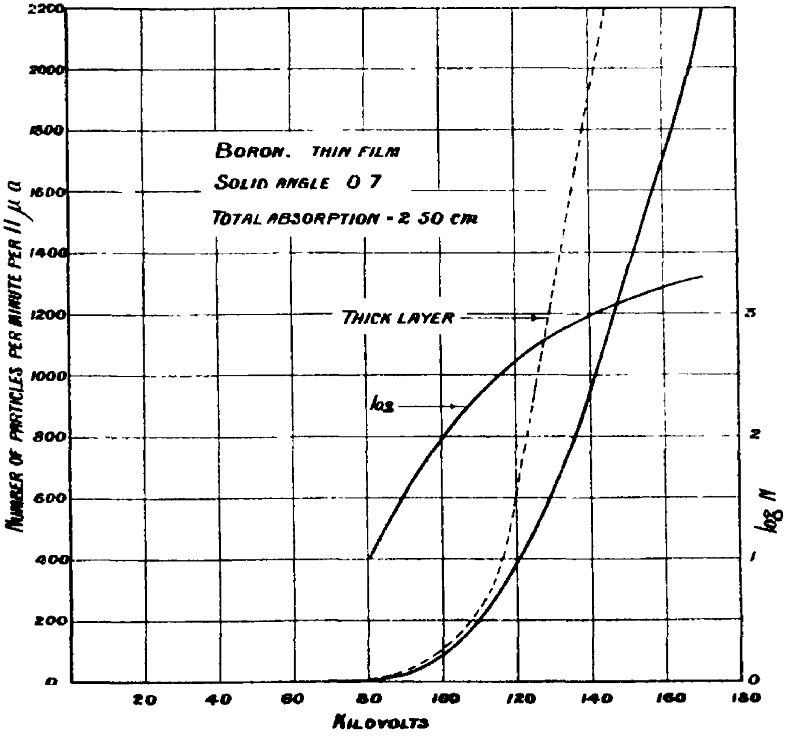

Boron. thin film
Solid angle 0 7
Total absorption - 2 50 cm
Thick layer
log
Number of particles per minute per 11 μ a
Kilovolts
log N
0
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
20
40
60
80
100
120
140
160
180
0
1
2
3

Fig. 5.

usual magnitudes we find that there is sufficient boron present to give a uniform layer of about 0·7 molecules thick. This could not have been deposited uniformly, but the film was certainly thin enough at all points for us to assume single collisions only with bombarding protons. It is seen that particles are detected at about 70 kv. and the numbers increase more rapidly with increase of bombarding energy than with the lithium film. The logarithmic plot shows

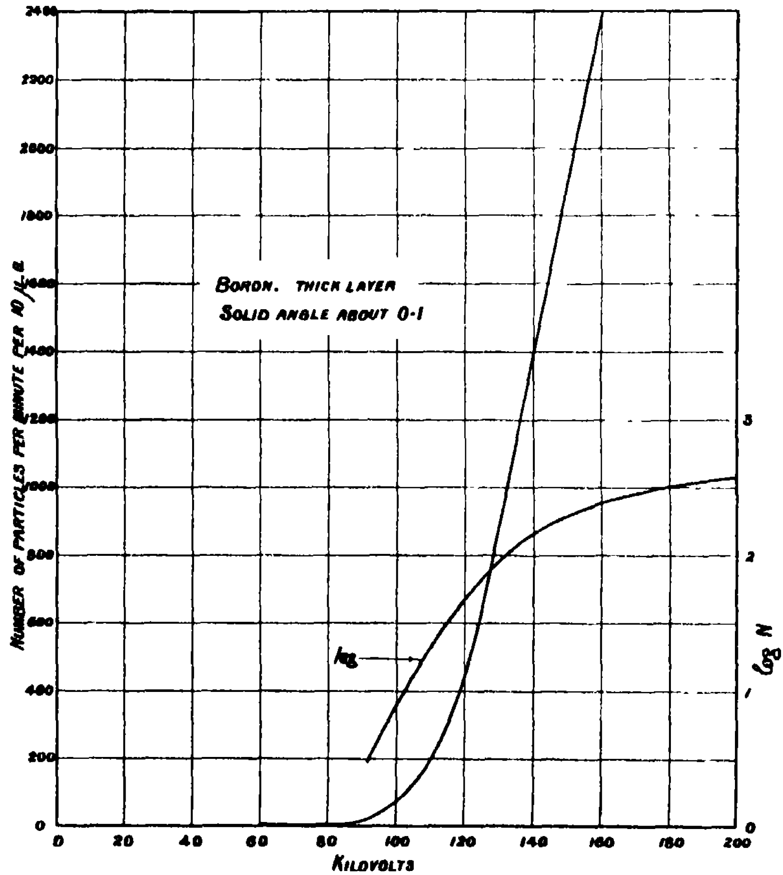


Fig. 6.

that in this case also we might expect the number of particles to become constant or even fall off above 200–300 kv. bombarding energy. The dotted curve in the same figure indicates the voltage variation with a thick layer of powdered boron element. This curve is given in greater detail in fig. 6. The number of particles found is larger than for the thin layer, but the form of the curve is very similar.

By an appropriate increase in the magnetic field it is possible to bring the H_2^+ ions on to the target and determine their effect separately, as shown in fig. 7. From the point of view of disruptive collisions the H_2^+ ion may be thought of as a pair of protons travelling together with an electron. The binding energy between them is negligible compared with the kinetic energy, which for either proton is one-half the energy of the particle. Hence a given current of molecular ions represents a current of protons of twice the magnitude, but with half the energy. We would therefore expect to begin to detect disintegration particles at about twice the energy found for the protons, and that the curve would rise twice as steeply. It is seen that these assumptions are approximately fulfilled.

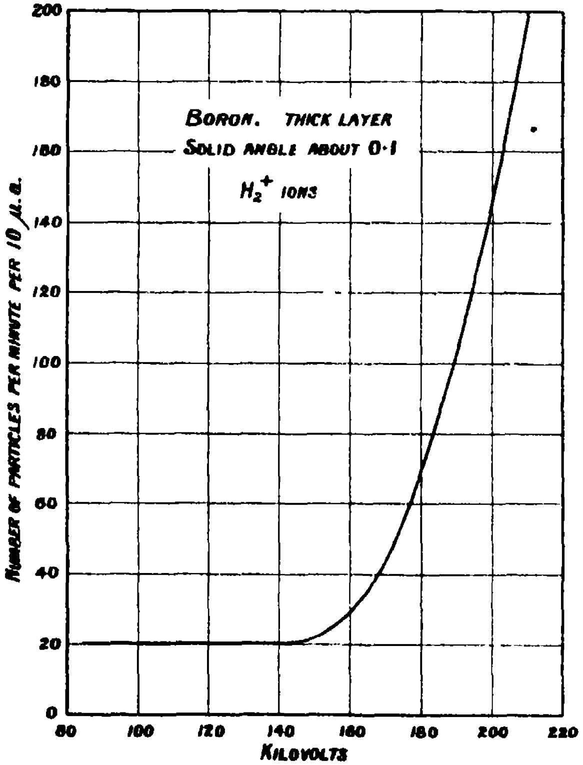


Fig. 7.

The peculiar steady emission of particles at low energies may be due to instrumental causes such as the presence in the beam of scattered protons, or it is just possible that there may be present in the molecular beam some component particle which is very efficient in causing disintegrations. If this should be the ion of the hydrogen isotope of mass 2 it could easily be increased in concentration and its effect studied in detail.

Range of the Particles from Boron.—Rough observations indicated that the particles emitted from boron by H^+ and H_2^+ were identical in range. Careful measurements were then made of the particles produced by protons, the observations being made in three sections. The first dealt with the end part of the range, beyond 2·5 cm., and is shown in fig. 8. It is seen that the number of particles falls off rapidly as the absorption is increased, but that there is a very definite end point. This maximum range has been determined with some care and appears to occur at very nearly 4·7 cm. The probable error in this quantity is estimated to be not greater than ± 1·5 mm., as it is obtained by direct comparison with the particles from polonium, which are assumed to have a range of 3·80 cm. at 15° C. and 760 mm. pressure. The rest of the absorption curve is shown, together with the end part on a very different scale, in fig. 9. The intermediate portions were obtained by the use of an absorption chamber 1 cm. deep, with windows 0·575 cm. equivalent absorption. The probable error over this part of the curve is estimated to be about ± 1·5 mm. The upper part of the curve was obtained with a canalized beam of particles by gradually moving the counting chamber away from a 0·575 cm. window by increments of known amount. In order to establish the form of this part of the curve more than 500,000 particles were counted in groups of about 1500. It must be remembered that the probability variation in the count at any point is compounded of the ordinary statistical variation together with varia-

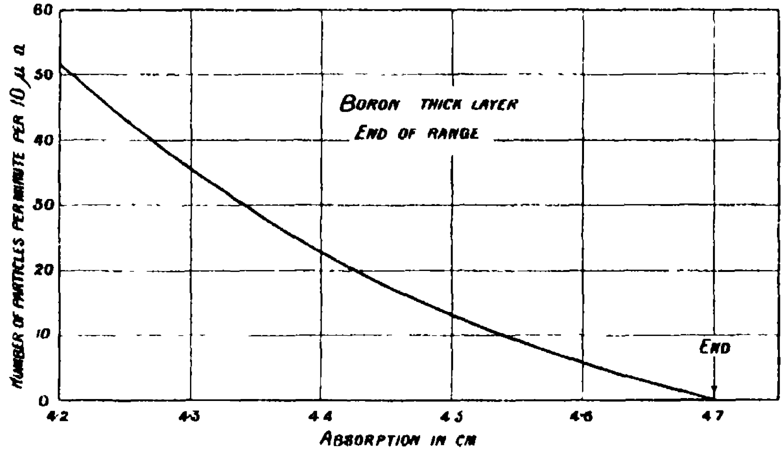


Fig. 8

tions due to changes in the accelerating field derived from the A.C. mains, and of the magnetic field which is bending the particles The part of the curve beyond 0·775 cm. is obtained by extrapolation in what appears to be an obvious manner. The ratio of the number of particles at 3·04 cm. to the number at 0·775 cm. absorption was obtained by making 25 counts at each of these points alternately without any change in the conditions The mean of these counts gave for this ratio 1·62%

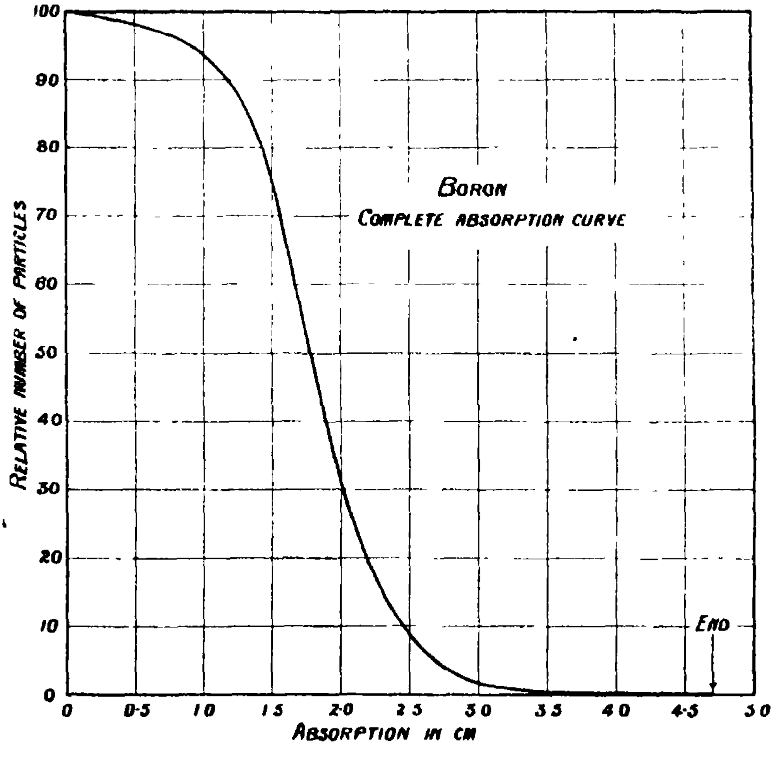


Fig. 9.

Light Elements Iron, Oxygen, Sodium, Aluminium, Nitrogen—In order that the above results might be assumed to be due to lithium and boron respectively and not appreciably to the iron target holder or to the oxygen or sodium present in the borax and oxides used, we made observations of the effect of bombarding these elements at energies up to 200,000 volts.

With a solid angle of collection of about 0·7, and a proton current of 20 microamperes, carefully cleaned iron or mild steel gave not more than two particles per minute, while on oxidizing a steel surface by heating it red hot in air no extra particles were observed. Commercial sodium metal gave only four particles a minute, while a sample of pure aluminium, kindly supplied by Professor Kapitza, gave the same number. We were unable to detect with certainty any effect from NH_4NO_3, so that it seems unlikely that nitrogen gives any large emission of particles over the range of proton energies investigated.

Beryllium and Fluorine.—A pure sample of beryllium oxide, kindly supplied by Dr. C. Desch of the National Physical Laboratory, was bombarded as strongly as possible with both protons and molecular ions. A small effect was observed with protons at the highest voltages, but the number of particles was too small to enable any definite measurements of range or of voltage variation to be obtained. It is therefore uncertain whether the particles arose from the beryllium itself or from a minute trace of boron or other impurity on the surface of the target.

Fluorine was bombarded as iron fluoride (Kahlbaum). Fig. 10 (*a*) shows the way in which the number of particles observed varies with the accelerating potential applied to the protons, while (*b*) is the absorption curve for the particles emitted at about 200 kv. Although the number of particles is too small for the detailed analysis of the range to be possible the course of the two curves renders it probable that the particles arise from fluorine itself and not from a contamination of boron.

Much higher accelerating potentials will have to be applied to the protons to enable any reliable information about the energies of the particles from Be and F to be obtained.

Heavy Elements—Gold, Lead, Bismuth, Thallium, Uranium, Thorium.—Cockcroft and Walton observed marked effects when they bombarded lead with protons, while under the same conditions the radioactive emission of α-particles from uranium oxide appeared to increase by an amount which did not depend markedly on the energy of the bombarding protons. This was a very remarkable observation and it is essential to test whether these elements can be disintegrated at low voltages and, if so, whether they exhibit any evidence of a " resonance " effect which might facilitate the passage of relatively low speed particles into favoured levels in the nucleus. Our first experiments confirmed the observations of Cockcroft and Walton, but we found it impossible to repeat the results from day to day. It was especially noticeable that the activity of the target was considerably increased after the accelerating tube had gone soft for a minute or two, as happens sometimes during a run. We therefore became suspicious that the effects might be due entirely to an extremely active deposit on the surface of the target. Measurements of the ranges of the disintegration particles showed that they were of about the same range as the particles from boron, and the fact that our accelerating tube was a large cylinder of pyrex, a borosilicate glass, strengthened the supposition that the whole effect might be due to boron Thallium and bismuth showed similar

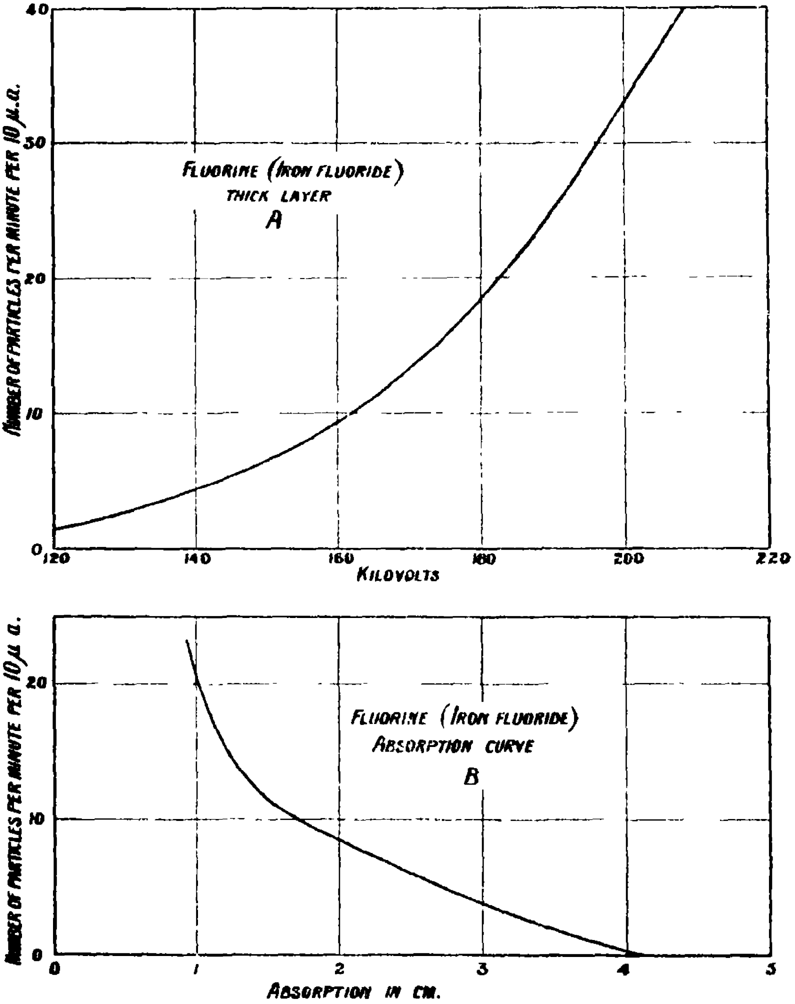


Fig. 10.

effects, while gold, which gave nothing at first, became active after the passage of a discharge through the accelerating space.

A cylinder of a different glass was now substituted for the pyrex and great care was taken to free the whole apparatus from all traces of boron or lithium. Fresh targets were prepared from the purest materials we could obtain, these being provided, for the most part, by Professor Kapitza. Under these conditions we detected no effect whatsoever from any of the elements mentioned at the head of the paragraph for proton energies up to 200 kv. These observations do not exclude the possibility of the emission of disintegration particles at higher energies, for boron only begins to show an effect at twice the energy required by lithium, and it is possible that the minimum energy required for a proton to penetrate a nucleus increases very rapidly with atomic number.

The Emission of Radiation.--Most substances fluoresce to some extent where struck by the beam of protons, but in addition the heavy elements show the emission of a form of radiation, probably in the long X-ray region, which is completely absorbed in a few centimetres of air and which ionizes gases with great efficiency. This radiation gives rise to an inconvenient background "wobble" in the output from the counting chamber, so that the oscillograph records of particles are superimposed upon a continuous oscillation. We have not looked with any care for evidence of the emission of γ-radiation

Discussion

A discussion of the disintegration function for thin films of lithium and boron will be postponed till results at much higher bombarding energies are available, and it can be said with certainty whether curves such as figs. 3 and 5 do show a maximum. However, the form of the curve calculated from Gamow's theory* is in general accord with the experimental facts if an appropriate cross-section is assumed for a collision.

The Disintegration of Boron. As pointed out by Cockcroft and Walton it is probable that the observed disintegration of boron is due mainly, if not entirely, to a splitting of the B_{11} nucleus into three α-particles when a proton penetrates the nuclear potential barrier and is captured. Measurements of the mass defect of the B_{11} isotope† indicate that such a transformation would lead to a release of about 11 million volts of energy. If each α-particle took one-third of this energy and no γ-ray were emitted the disintegration particles would possess a single definite energy of about 3·7 million volts, corresponding with a range in air of approximately 2·4 cm. The experimental curve given in fig. 9 which was obtained using a narrow beam, shows no evidence whatsoever of a single range of this magnitude. By substituting a weak source of polonium for the target used in the experiments the form of the absorption curve for a homogeneous group of α-particles can be determined under our experimental conditions. It was found that the number of such particles counted remained constant to within 3 mm. of the end of the range, and showed an almost linear drop to zero over the final 1·5 mm. A slight "tail" rounded off the end of the curve. The experimental curve for the disintegration particles from boron shows no evidence of any homogeneous group of particles with a definite range when examined on this criterion. We are forced to the conclusion that there are present particles of all possible velocities, from zero to the maximum.

---

* "Constitution of Atomic Nuclei and Radioactivity," 1931.

† Aston, 'Proc. Roy. Soc.,' A, vol. 215, p. 487 (1927); Bainbridge, 'Phys. Rev.,' vol. 63, p. 424 (1933).

Without making any particular assumption about the actual process of capture and disintegration of the B nucleus we might anticipate that the most probable mode of disintegration would be for the three α-particles to escape symmetrically with equal velocities, making an angle of 120° with one another, fig. 11 (*a*), the probability of other modes of disintegration diminishing rapidly with departure from the symmetrical distribution

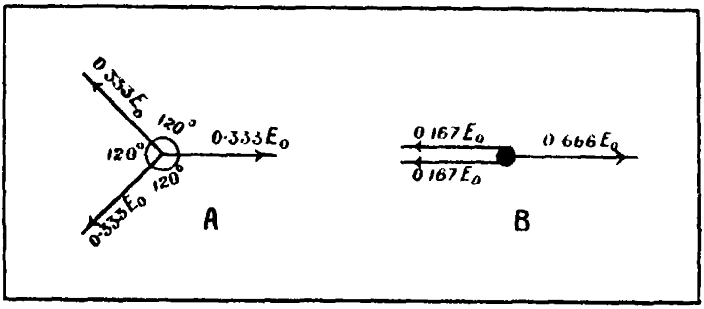


FIG. 11.

In our experiments the α-particles which were counted escaped from the target at right angles to the direction of the incident beam of protons. It is natural to assume that there is no favoured direction of emission* but that on the average the α-particles are liberated equally in all directions. To simplify the calculations we assume that two of the α-particles are emitted symmetrically with regard to the third particle, and thus have equal velocities. If u_1 is the velocity of each of the symmetrical particles, θ the angle between them, and u_2 the velocity of the third particle, fig. 12, then from the equivalence of momentum

$$2u_1^2(1+\cos\theta) = u_2^2. \tag{1}$$

---

* For simplicity in the subsequent calculations the momentum of the incident proton is neglected, since in general it is small compared with that of the released α-particles.

If we assume the same total energy $\frac{1}{2}mu_0^2$ (where m is the mass of the α-particle and u_0 the equivalent velocity) is released in all disintegrations,

$$2u_1^2 + u_2^2 = u_0^2. \tag{2}$$

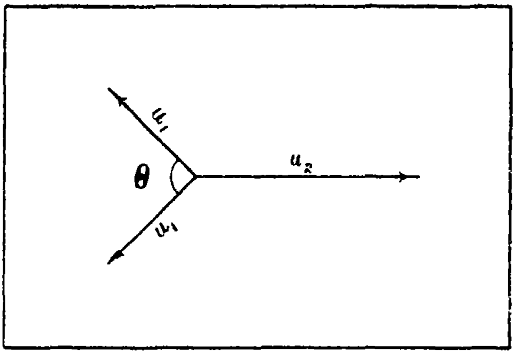


Fig. 12.

From (1) we obtain by substitution

$$\left.\begin{aligned} u_1^2/u_0^2 &= \frac{1}{4 + 2\cos\theta} \\ u_2^2/u_0^2 &= \frac{1 + \cos\theta}{2 + \cos\theta} \end{aligned}\right\}. \tag{3}$$

The value of u_2^2/u_0^2 is a maximum and equal to 0·666 when $\theta = 0$, *i.e.*, when one α-particle is emitted at 180 with the pair of particles of equal velocity, fig. 11 (*b*) The value of $u_1^2/u_0^2 = 0{\cdot}50$ is a maximum when $\theta = 180°$ and a minimum, 0·167, when $\theta = 0$.

Since the maximum energy acquired by an α-particle is 0·666 E_0 this energy is taken to correspond with that of the particle of maximum range 4·7 cm. observed in our experiments, *i e.*, to 5·96 million volts. The total energy appearing in the disintegration is thus 8·94 million volts. The energy of each of the α-particles in the most probable mode of emission at 120° would be 2·98 million volts, corresponding with a range of about 1·8 cm. in air. Reference to the experimental curve, fig. 9, shows that at this range the slope of the curve is a maximum, indicating that this is actually the most probable range. This agreement gives strong support to our general hypothesis.

From equation (3) energy curves are drawn in fig. 13, showing (*a*) the way in which the energy of any one of the pair of particles varies with the angle θ, and (*b*) the corresponding variation in energy of the single opposing particle. It is evident that only the latter single particle can contribute to particles of energy less than 0·167 E_0 or more than 0·50 E_0, while intermediate values are made up from both groups of particles.

On these views the " tail " of the absorption curve, fig. 8, is due to the distribution of particles of energy between 0·50 E_0 and 0·666 E_0, *i.e.*, to values of θ between $\theta = 90°$ and $\theta = 0°$, corresponding to ranges between 3·0 cm. and 4·7 cm. in our experiments. It is seen that in this part of the curve the numbers of particles fall off rapidly with increasing energy, *i.e.*, with decrease of θ between 90° and 0°.

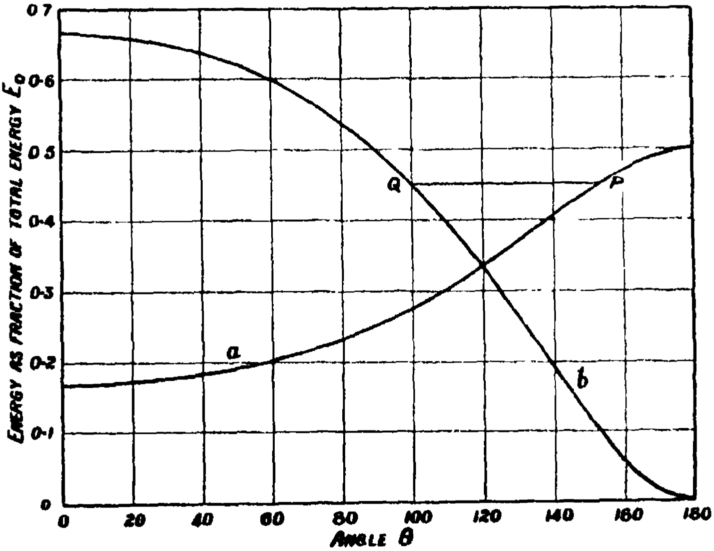


FIG. 13.

This at once gives us important information about the way the probability of emission varies with angle θ between 0° and 90°. If we take the experimental curve of fig. 8 connecting numbers and range between 3·04 and 4·7 cm., and plot the corresponding curve between number and angle, the numbers are found to fall nearly exponentially with θ, decreasing to half value in about 10°. It is of great interest to see whether this simple expression for the change of probability of emission with angle holds generally. This can be tested by calculating the absorption curve of the α-particles on this hypothesis and comparing it with the observed curve. For this calculation we suppose that the number of particles emitted corresponding to any angle θ falls off exponentially with the difference between this angle and 120°, *i.e.*,

$$N_{\theta(d\theta)} \propto e^{-\alpha\phi}, \tag{4}$$

where $\phi = 120° - \theta$ for θ less than 120°, and $\phi = \theta - 120°$ for θ greater than 120°.

The number of particles emitted with energies greater than that corresponding with any given range can then be calculated by graphical integration as follows

From the curves given in fig. 13 the angles which contribute particles of a definite energy may be read off. For example, consider the energy corresponding with the horizontal line PQ, fig. 13. Two α-particles have energy corresponding to P in the (*a*) curve and one to Q on the (*b*) curve, and the corresponding angles can be read off If the energy corresponding to the line PQ represents the energy lost in the absorber, the number penetrating the absorber and finally collected and counted will be the number emitted over the angular range from P to the angle giving the maximum energy on the (*a*) curve, *i.e.*, from θ_P to $\theta = 180°$, and over the range from θ_Q to $\theta = 0°$ on the (*b*) curve. The distribution function (4) may now be integrated over these ranges, and the total number of particles emitted thereby found. This was done for representative points all along the absorption curve. In order to determine the value of α more accurately than was possible from the end part of the absorption curve alone, use was made of the ratio of the numbers of particles collected at 0·8 and at 3·04 cm. absorption, which ratio had been determined accurately for this purpose as 1·62% The numbers were calculated for the two absorptions by the method we have described and the calculated ratio equated to 1·62% In this way we found $\alpha = 0{\cdot}069$, in good agreement with the value found from the end part of the absorption curve The energy absorption curve calculated using this value of α is shown in fig. 14 This can be transposed readily to a range absorption curve by the application of Geiger's law connecting range and energy, together with the corrections to this law which have been determined experimentally.* This curve is also given in the figure. The circles in this last curve are points from the experimental curve given in fig. 9.

Considering the simplicity of the assumptions the agreement between calculation and experiment is remarkably good, and indeed better than would be expected. Not only does the theory account for the main features of the distribution of α-particles with range, but it is in excellent numerical agreement throughout the whole curve, where the number of particles for different ranges varies more than 1000 times. While the exponential law of distribution assumed accords well with the "tail" part of the absorption curve, it should be pointed out that the main part of the curve is comparatively insensitive to variation in the law of distribution in the neighbourhood of 120°. In order to make the calculations possible we had to assume that two of the particles had identical velocities When the difference between the angle of emission and 120° is small this may not be so, but over this range the form of the curve is relatively insensitive, as we have seen, to the exact assumptions made. When the angle of emission differs widely from 120° i.e., for the beginning and end of the absorption curve, two of the particles must possess very nearly equal velocities in order to satisfy the momentum laws. The simplifying assumption is therefore justified, and in any case it seems likely that the main features of the absorption curve would not be seriously altered if it were possible to calculate it on the assumption that all three particles could possess different velocities.

---

* *Cf.* Rutherford, Chadwick, and Ellis, "Radiations from Radioactive Substances," p. 102 *et seq.* (1930).

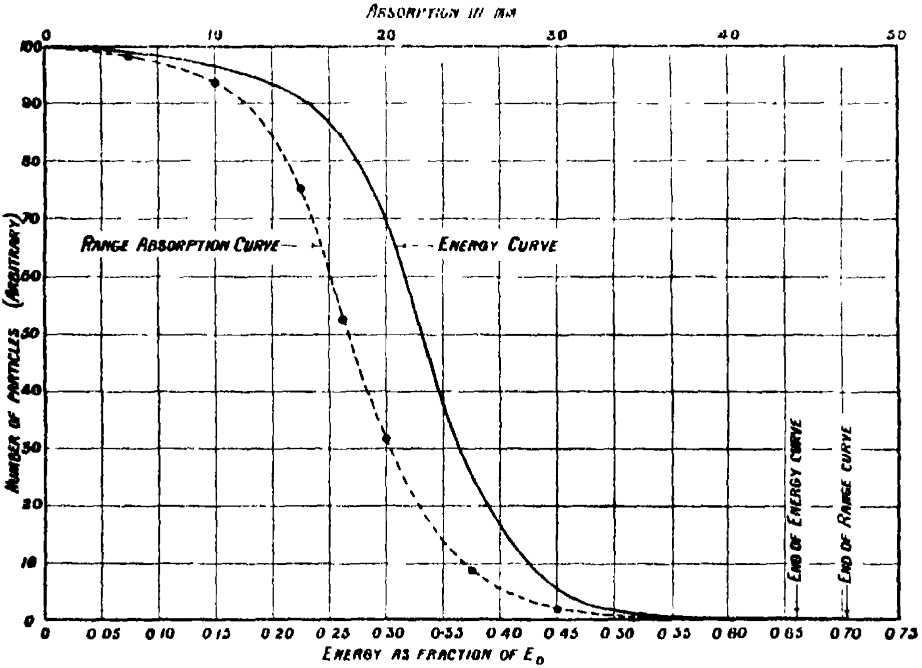


FIG. 14.

On the views we have outlined the total energy released in the boron disintegration is 9 million volts, while the available energy calculated from the mass is somewhat greater, viz., about 11 million volts. We have not so far had time to examine whether appreciable γ-radiation arises from the bombardment of boron. If the emission of high frequency γ-radiation should accompany the majority of these transformations, it is obvious that the simple theory we have outlined would no longer be valid. We hope to examine this point subsequently.

It is obviously of great importance to obtain more direct evidence of the mode of disintegration of boron. We have constructed a special counting chamber and hope soon to test whether three α-particles are liberated simultaneously. A still more direct method of settling this important point is to examine the tracks of the emitted α-particles in an expansion chamber. Experiments in this way have already been made by Kirchner in Munich and by Dee in this laboratory. Decisive experiments by this method are much more difficult to carry out for practical reasons for boron than for lithium, on account of the relative shortness of the range of the α-particles emitted in the transformation of boron. We have already seen that in the symmetrical case of emission the range of each of the three α-particles is only 1·8 cm. It will, however, be of the greatest importance to determine the distribution of the particles both as regards angle and range

We desire to express our thanks to Mr. R. B. Kinsey, who has helped in many of the experiments and in the development of the high tension voltmeter; and to Drs Cockcroft and Walton for much useful information, and to G. R. Crowe for his assistance throughout the investigation.

Summary.

(1) An apparatus is described which enables an intense beam of hydrogen canal-rays to be produced and then accelerated with potentials up to 200,000 volts. The different kinds of ion present in the beam are separated by bending in a magnetic field, and any chosen constituent can be allowed to fall upon a target of the material to be examined. A mica window subtending a solid angle of about 0·7 allows any disintegration particles produced to pass into a counting chamber. In this way it is possible to examine the disintegration of some elements at comparatively low voltages. The voltmeter used for measuring the high D.C. potentials is described, and methods used for measuring the ranges of the particles are indicated.

(2) Lithium gave particles at 30 kv. with protons bombarding the target, while the effect from boron was not appreciable below 60 kv. The voltage variation of the yield of disintegration particles was measured for very thin

films of Li and B, as well as for thick films. Molecular hydrogen ions gave particles at approximately twice the energy required for protons and the yield increased at twice the rate as the energy was increased.

(3) Careful measurements of the range of the particles from boron are described and it is shown that there is a maximum range of 4·7 cm. By far the greater number of particles possess ranges distributed about 1·8 cm., but there are present particles of all ranges from zero up to the maximum.

(4) It was shown that Fe, O, Na, Al, N gave no detectable effect when bombarded with protons or molecular ions at energies up to 200 kv.

(5) Be and F gave detectable effects, but the yield was too small to enable accurate measurements to be made.

(6) Au, Pb, Bi, Tl, U, Th gave no observable effect at our bombarding energies. Slight traces of active impurities, probably boron, gave large effects until they were eliminated by suitable means.

(7) A discussion is given of the disintegration of the B_{11} nucleus by proton capture. It is assumed that the nucleus disintegrates into three α-particles which escape with a probability distribution of directions about the symmetrical coplanar 120° position, and it is shown that this distribution may be obtained and the absorption curve calculated.

(8) The total energy produced in a disintegration on the theory we have described is about 9 million volts, whereas the energy available calculated from the mass of the B_{11} nucleus and the disintegration particles is 11 million volts.

Results of Calculations of Atomic Wave Functions. I.—Survey, and Self-consistent Fields for Cl^- *and* Cu^+.

By D. R. Hartree, F.R.S., Department of Mathematics, Manchester University.

(Received May 1, 1933.)

1. *Introduction.*

An approximation to the structure of a many-electron atom can be obtained by considering each electron to be a stationary state in the field of the nucleus and the Schrodinger charge distribution of the other electrons, and rather more than five years ago I gave a method of working out atomic structures based on this idea, and called the field of the nucleus and distribution of charge so obtained the "self-consistent field."*

The method of working out the self-consistent field for any particular atom involves essentially (*a*) the estimation of the contributions to the field from the various electron groups constituting the atom in question; (*b*) the solution of the radial wave equation for an electron in the field of the nucleus and other electrons, this solution being carried out for each of the wave functions supposed occupied by electrons in the atomic state considered; and (*c*) the calculation of the contribution to the field from the Schrodinger charge distribution of an electron group with each radial wave function. The estimates of the contributions to the field have to be adjusted by trial until the agreement between the contributions finally calculated and those estimated is considered satisfactory.

In the paper in which the method was first suggested, some results were also given, and since then, results for a number of atoms have been worked out by myself and others, and some of these have been published. But the results published have usually been results for the whole atom, such as the total charge distribution or field; detailed results such as individual wave functions, and contributions to the field from the different electron groups have not been given except for O, O^+, O^{++}, O^{+++}, and Si^{+4}.† These individual wave functions and contributions to the field, however, are often the quantities required in applications. A considerable amount of information of this kind is available, and has been supplied privately to anyone who knew of its existence and asked for it, but for some time it has seemed desirable to publish it so as to make it more widely available.

---

* Hartree, 'Proc. Camb Phil. Soc,' vol. 24, pp. 89, 111 (1928).

† Oxygen: Hartree and Black, 'Proc. Roy. Soc.,' vol. 139, p. 311 (1933). Silicon McDougall, 'Proc. Roy. Soc.,' vol. 138, p. 550 (1932).

There were, however, some objections to publishing it as it stood. As experience of working out atomic structures has increased, several improvements have been made in numerical technique, and also the standard of what is considered a "satisfactory" agreement between estimated and calculated contributions to the field has gone up considerably. A good deal of the available information was based on old work in which neither the degree of agreement between estimated and calculated fields, nor the numerical accuracy to which the work was carried out, now seem adequate. It seemed desirable therefore to carry out a revision of the work, using the improved methods and working to a higher numerical accuracy (principally in order to lessen the effect of possible cumulative integration errors), and also aiming at a higher standard of agreement between estimated and calculated contributions to the field, before publishing the detailed results. That such revision is desirable is shown by appreciable differences in some cases between the revised and earlier results.

The method of carrying out the numerical work is now fairly well standardized, and the revision of the calculations is largely a matter of expert computing. A grant was made by the Government Grants Committee of the Royal Society for the purpose of employing the professional assistance of expert computers in connection with this work, and I wish to acknowledge my thanks to the Committee for its assistance, to Miss D. S. Greene,* for undertaking the arrangements for getting the work done, and to Dr. L. J. Comrie and Mr. D. S. Sadler who actually did the calculations, for the very satisfactory way they have carried out the computing work which is of rather unusual and not altogether straightforward kind.

With this assistance available, the work of improving the earlier results of any particular atom is being carried out as follows. From the earlier work, I make revised estimates of the contributions to the field from the various electron groups, and the computing work carried out professionally is concerned with the calculation of wave functions in the field so constructed, regarded as *given*, and of charge distributions from these wave functions. For reference I will call these calculations the "standard calculations." Unless estimates of the contributions to the field have been unusually fortunate, the results of these standard calculations are not yet near enough to the self-consistent field to be quite satisfactory, but they should be near enough for the effect of any variation of the estimates to be treated as a first order variation from the results of the standard calculations.

---

* Director, Calculating and Statistical Service, Victoria House, Vernon Place, London, W.C.1.

A further revision of the estimates is made if necessary and the *variations* in the wave functions, etc., due to the *variations* in the estimates from those used in the standard calculations, are calculated, and the variations of wave functions, etc., added to the results of the standard calculations; the variations are so small that this variation calculation is very much shorter and easier than the main calculation. If necessary, further revisions of the estimates are made and corresponding variations from the results of the standard calculations are worked out, until a satisfactory approximation to the self-consistent field is obtained. For the atoms for which results are given here, I have myself been responsible for carrying out these final stages of the work.

This paper presents the results of this work for two atoms, Cl$^-$ and Cu$^+$; the earlier work for both of these atoms was done by myself, and some results for Cl$^-$ were published in my first paper. Some results of the earlier work on Cu$^+$ have been published by Slater in connection with his approximations to the wave functions by analytical formulæ.* Other atoms for which such work is in progress are Al^{+3}, K$^+$, Rb$^+$, Cs$^+$.

The results for Cu$^+$ here given are interesting, as they are the first which have been obtained for an atom with an outer shell of 18 electrons; the other atoms for which calculations have so far been carried out are atoms for which the outer shell is a complete or incomplete 8-shell. Comparison of the results for Cu$^+$ and the alkali metal ions may throw light on the characteristic difference of properties of the two sub-groups of the first column of the periodic table.

Since the method of obtaining approximate wave functions for a many-electron atom by means of the "self-consistent field" was suggested, a better approximation has been indicated independently by Slater† and Fock‡, and the equations obtained by the latter; but the numerical application of Fock's approximation involves difficult problems of numerical technique for any but the lightest atoms, and as far as I know, no case has yet been worked out quantitatively. Slater and Fock have shown that the approximation made in the "self-consistent field" method is the best for its simplicity that could be obtained; also it is probable that in the solution of Fock's equations for any atom the first step would be the approximate solution of the equations of the self-consistent field, so that as well as being the best results at present available, the results, of which those presented here form a part, are also steps on the way to better approximations.

---

* Slater, 'Phys. Rev.,' vol. 42, p. 33 (1932).

† 'Phys. Rev.,' vol. 35, p. 210 (1929).

‡ 'Z. Physik,' vol. 61, p. 126 (1930).

§ 2. *Summary of Notation and Nomenclature.*

All quantities are supposed measured in natural atomic units, that is units such that the measures of the mass of the electron, of the magnitude of the charge on the electron, and of $h/2\pi$, are all 1; the unit of length is

$$h^2/4\pi^2me^2 = 0\cdot532 \text{ A.U}$$

If, in spherical polar co-ordinates with origin at the nucleus, the wave function of a single electron in a central field of potential v is written

$$\psi(r, \theta, \phi) = \frac{P(r)}{r} S_l(\theta, p), \tag{1}$$

where S_l is a spherical harmonic of degree l, $P(r)$ is called the radial wave function, and satisfies the equation*

$$d^2P/dr^2 + [2v - \varepsilon - l(l+1)/r^2] P = 0, \tag{2}$$

where ε is the negative energy parameter reckoned with the ionization energy of the hydrogen atom as unit, so that if ν is the corresponding wave number, $\varepsilon = \nu/R$. It is sometimes convenient to write the quantum numbers n, l of the wave function of which $P(r)/r$ is the radial part as suffixes [viz., $P_{n,l}(r)$], or alternatively in the form $P(nl|r)$; and for specifying the value of l it is best to adopt the usual conventional letter according to the scheme

	s	*p*	*d*	*f*	*g*	*h*	.
for $l =$	0	1	2	3	4	5	...

If $P(r)$ is normalized so that $\int_0^\infty P^2\,dr = 1$, then $P^2\,dr$ is the charge lying in a spherical shell between radii r and $r + dr$, on Schrodinger's interpretation of the wave function, and P^2 may conveniently be called the "radial density" of the charge distribution of the electron concerned; if P is not normalized the radial density is $P^2 \Big/ \int_0^\infty P^2\,dr$.

---

* *Cf* Hartree, *loc. cit.*

If we write*

$$Z_0\,(nl,\,nl\,|\,r) = \int_0^r [\mathrm{P}\,(nl\,|\,r)]^2\,dr \bigg/ \int_0^\infty [\mathrm{P}\,(nl\,|\,r)]^2\,dr, \tag{3}$$

then $Z_0(nl,\,nl\,|\,r)$ is the total charge lying inside radius r, for an electron in an $(n,\,l)$ wave function. If the field of the atom is considered to be spherically symmetrical, and Z is the "effective nuclear charge" at radius r, defined so that the field† at radius r is Z/r^2, then for an atom of nuclear charge N

$$Z = N - \Sigma\, Z_0\,(nl,\,nl\,|\,r)$$

the sum being over all occupied wave functions There may be a number of wave functions with the same values of (nl); if there are N_{nl} wave functions‡ with given values of n and l, then

$$Z = N - \Sigma_{nl} N_{nl} Z_0\,(nl,\,nl\,|\,r) \tag{4}$$

the sum being now over all values of n and l; if the (nl) group is complete, $N_{nl} = 2\,(2l + 1)$.

If C is the total charge on the atom,

$$C = N - \Sigma_{nl} N_{nl},$$

so (4) can be written alternatively

$$Z = C + \Sigma_{nl} N_{nl}\,[1 - Z_0\,(nl,\,nl\,|\,r)]. \tag{5}$$

For an electron occupying an (nl) wave function, $[1 - Z_0\,(nl,\,nl\,|\,r)]$ is the charge lying outside radius r, and tends to 0 for large r; $[1 - Z_0\,(nl,\,nl\,|\,r)]$ may be described as the contribution to Z at radius r from an electron occupying an (nl) wave function, and unit charge on the nucleus, it is in terms of these contributions to Z that work on the self-consistent field is usually done, and the extent of the agreement between estimated and calculated contributions

---

* This notation is unnecessarily elaborate for consideration of the self-consistent field alone, but is used for consistency with the notation which is necessary in some applications of the results (see Hartree and Black, *loc. cit.*), and which is also convenient for the extension to Fock's equations.

† It is important to remember that when Z varies with r, the potential at radius r is not even approximately Z/r, in general.

‡ If the (nl) group is not complete, its contribution to the field is not spherically symmetrical and for the purpose of the self-consistent field we take a spherical average. Slater has shown that the approximation involved in taking this average is of the same order as that involved in omitting the interchange terms which form the difference between Fock's equations and the equations of the self-consistent field.

In the results of which this paper gives a first instalment, the agreement will be considered "satisfactory" if the difference between the estimated and calculated contributions $N_{nl}\ [1 - Z_0\,(nl, nl\,|\,r)]$ to Z from each whole group does not exceed 0·02 at any radius, and the difference between the sums of the contributions for the whole atom does not exceed 0·03 at any radius, and probably the actual results presented will usually be well within these limits.

For the standard calculations and subsequent work based on them, the contributions to Z have been estimated to 0·005, and the calculated values have been calculated to three decimals, the last not always certain to 1 or 2. For all but the outermost groups, it has been quite practicable to estimate the contributions to Z so that the difference between estimated and calculated values is nowhere greater than 0·005, and for the outermost group it has been possible to attain a maximum difference of 0·01 or only slightly over. The atoms for which results are given here are the two most troublesome of all for which calculations have been carried out, on account in each case of the sensitiveness of the outer groups to the estimated contributions to the field, and the order of agreement of estimated and calculated contributions to Z, and so of approximation to the self-consistent field, attained for them should always be attainable for others with less trouble.

The contributions to Z may be called "stable," in the sense that if the estimated contributions from any group is increased over a range of r, the effect of this is to decrease the calculated contribution from this group (and from others also); for an increase of Z means that the attractive field on an electron towards the nucleus is increased, the wave functions of electrons in the field become more compact, and the proportion of the electron distribution lying inside any given radius is increased; that is Z_0 for that group is increased, and $1 - Z_0$ is decreased; and $1 - Z_0$ occurs with a positive sign in (S).

For all but the groups of the outermost shell, it is usually if not always the case that the change in the calculated contributions is smaller than the change in the estimated contributions. If this were so for all groups, an iterative process, taking the calculated contributions from one approximation as the estimates for the next, would give a series of calculations with results converging to those for the self-consistent field; though this process would be unnecessarily lengthy, as with experience it is usually possible to make revised estimates better than those obtained by simply taking the calculated contributions of the previous approximation. But for the groups of the outer shell, and particularly the most loosely bound group, there sometimes occurs a phenomena which may be termed over-stability, in which a change of estimated contribution to Z causes a change in the calculated contribution larger than a change in the estimate. When a group is over-stable in this sense, an iterative process would not converge, but, for small variations, would oscilllate and diverge, and it is then quite necessary to choose, as revised estimates of contributions to Z, values better than the calculated contributions of the previous approximation; it is also unusually difficult to make satisfactory estimates and adjustments to them, so the process of approximation to the self-consistent field is most troublesome in such cases.

A numerical example of over-stability is given in the discussion of the results for Cu^+ in § 4.

§ 3. *Description and Explanation of the Tables.*

It is proposed to give the following results for each atom for which calculations on these lines are carried out:—

(*a*) A table of the radial wave functions $P(nl|r)$, not normalized, and also P/r^{l+1} for small r.

(*b*) A subsidiary table giving for each wave function the value of the energy parameter ε in the radial wave equation (2), and the value of the normalizing integrals $\int_0^\infty P^2\,dr$.

(*c*) A table of the contributions $2\,(2l+1)\,[1-Z_0\,(nl,\,nl\,|\,r)]$ to Z at radius r. [The atoms for which calculations are at present in progress all consist of complete groups so that $N_{nl}=2\,(2l+1)$.]

(*a*) *and* (*b*).—The radial wave functions are tabulated unnormalized for several reasons.

In applications of the results, integrals of the form $\int_0^\infty P\,(\alpha\,|\,r)\,F\,(r)\,P\,(\beta\,|\,r)\,dr$, with normalized wave functions, often occur; but in evaluating them it is at least as easy to use unnormalized wave functions and divide by

$$\left[\int_0^\infty P^2\,(\alpha\,|\,r)\,dr\right]^{\frac{1}{2}}\left[\int_0^\infty P^2\,(\beta\,|\,r)\,dr\right]^{\frac{1}{2}}$$

at the end, as to use normalized wave functions throughout. The values of the normalization integrals $\int_0^\infty P^2\,dr$ are given in the subsidiary table.

Also as $r \to 0$, P/r^{l+1} tends to a finite and non-zero value, and for numerical reasons it is convenient in the integration of the wave equation to keep this limit the same for different wave functions with the same l. When this is done, these different wave functions are nearly the same for r small, and advantage can often be taken of this in applications; this advantage is lost when normalized wave functions are used.

Also in dividing each calculated value of P by the square root of the normalizing integral, and rounding off to the number of places retained in the table, there is some loss of accuracy and the resulting rounded-off values may not be quite normalized.

It has therefore seemed best to tabulate actual calculated values of P in most cases. In almost all cases the tabulated values are rounded off from the values obtained on the calculation sheets, on which usually one, and occasionally two, more significant figures appear than in the tables; so that it is hoped that the divergence from the correct solution of the equations with the field actually used in the final calculations is generally not more than 1 in the last figure tabulated The field actually used was not of course exactly the self-consistent field, and it is rather difficult to estimate the variation in the calculated values which would result from replacing the actual field used by the self-consistent field, but in the results here given this variation would probably not be more than 1 or 2, in the last figure tabulated, for a wave function of an inner group, and might be rather more for a wave function of one of the outer groups (especially for $(3p)$ in Cl^- and $(3d)$ in Cu^+) but should not usually be more than 5

The calculations generally provide values of P at about twice as many values of r as those shown in the table, the intervals used in the numerical integration being usually half those used in the tabulation; the tables here given should be adequate for all ordinary applications, but the intermediate values of P can be supplied as required.

For some applications, P may be required for small r to greater accuracy, or at smaller intervals of r, than it is given in the main table, for such purpose it is best to work in terms of P/r^{l+1}, as this is finite and non-zero at $r = 0$, and is nearly linear in r for r small, and so is easy to interpolate, and a small table of this function for small r is given at the head of the main table.

A table is also given of the energy parameters ε for the different wave functions. Each of these is formally the negative energy of an electron in a wave function in the field of the nucleus and the rest of the atom, or the energy required to remove this electron from the corresponding wave function to rest at infinity, the rest of the atom being considered as providing a static field; this is not the same as the X-ray ionization energy of the atom, for the rest of the atom is not a static field but an electronic system whose configuration and energy changes when one electron is removed. But empirically there is a close quantitative correspondence between these ε's and the X-ray ionization energies, which has not been fully explained as far as I am aware.

(*c*) For each atom the contributions to Z are tabulated, partly because they will be useful as a starting-point for a better approximation to the self-consistent field if one is ever required, and also for the first steps in the solution of Fock's equations,* but more particularly because they are probably the most useful quantities in terms of which to carry out the interpolation between different atoms. Slater has suggested a method of interpolation of wave functions based on the approximate representation of the wave functions by analytical formulæ, and interpolation of the constants in these formulæ, but hardly enough atoms have been yet worked out to provide enough values of the constants between which to interpolate accurately; the change of atomic structure with atomic number is rather irregular (*cf.* the occurrence of "transition groups") so that the relations between the atomic number and the constants defining a wave function would not be expected to be very easy functions to interpolate accurately If fairly accurate wave functions are required, it seems best at present not to try to interpolate the wave functions themselves, but to interpolate the contributions to Z, which are probably much less sensitive to details of atomic structure, and calculate wave functions in the field so constructed. This must in any case be the process in beginning the calculation of the self-consistent field for a new atom, and it is mainly for this purpose that the contributions to Z are tabulated.

The contribution from each (*nl*) group is tabulated, not the contribution from one electron, as these contributions from each whole group are the quantities usually required. Three decimals are given, as although only two decimals and sometimes a 5 in the third place were used in the estimates, the variations of the calculated values with the variations in the estimates show that the third decimal in the calculated values is probably within 2 or 3 of its value for the self-consistent field in most cases, and almost certainly for the inner groups; so that the third decimal is of some value, although not certain to a unit.

---

* The functions occurring in Fock's equations, corresponding to the P's in the equations of the self-consistent field, are orthogonal functions, which the solutions of the self-consistent field equations are not; but the changes in contributions to Z due to making the P's orthogonal are probably smaller than those due to the resonance terms in Fock's equations, and as a first step the $Z_0(\alpha\alpha|r)$ functions which occur in Fock's equations can be taken as those of the self-consistent field.

§ 4. *Results for* Cl^-.

The maximum difference between estimated and calculated contributions to Z from any group, for the standard calculations, was 0·03; revised estimates, were made for all groups on the basis of the results of the standard calculations, and a second set of revised estimates was made for the $(3p)^6$ group only, and was used for the calculations giving the results tabulated. The maximum difference between estimated and calculated contributions to Z for the field used to give these results is less than 0·01 for each group, and the difference between estimated and calculated Z for the whole atom is also nowhere larger than 0·01; for the earlier calculations of which some results were given in my first paper, the maximum difference in the total Z was 0·08.

The wave function for the outermost (nl) group of a negative ion is very sensitive to the estimated field. Analytically this arises from the smallness of the energy parameter ε for such a wave function; it is related to the looseness of binding of this electron, although, as already emphasized, ε cannot be taken directly as a measure of the ionization energy. For large r, the behaviour of the wave function is chiefly determined by the value of ε; when the estimated field is changed, then ε for each wave function is changed, and if ε is small, a small change in ε may be a considerable *proportional* change and so affect considerably the behaviour of the wave function for large r. The $(3p)^6$ group is in fact over-stable in the sense explained at the end of § 2, and it is satisfactory that such a good agreement between estimated and calculated contributions was reached at the second revision of the estimates used in the standard calculations.

Table I.—Cl^- Wave Functions.

Tables of P/r^{l+1} for small r.

r.	$(1s)$.	$(2s)$.	$(2p)$	$(2s)$.	$(3p)$.
0·000	100	100	400	100	400
0·005	91·86	91·79	383·4	91·79	383·4
0·010	84·38	84·14	367·6	84·12	367·5
0·015	77·52	77·01	352·5	76·97	352·2
0·020	71·23	70·36	338·2	70·30	337·6

Table of P.

r	(1*s*)	(2*s*)	(2*p*).	(3*s*).	(3*p*).
0·000	0·000	0·000	0·000	0·000	0·000
0·005	0·459	0·459	0·010	0·459	0·010
0·010	0·844	0·841	0·037	0·841	0·037
0·015	1·163	1·155	0·079	1·155	0·079
0·020	1·425	1·407	0·135	1·406	0·135
0·03	1·805	1·752	0·280	1·749	0·280
0·04	2·034	1·921	0·460	1·913	0·458
0·05	2·150	1·951	0·664	1·937	0·661
0·06	2·183	1·872	0·884	1·850	0·878
0·07	2·156	1·709	1·113	1·678	1·105
0·08	2·087	1·483	1·346	1·441	1·335
0·09	1·989	1·210	1·579	1·158	1·563
0·10	1·873	0·906	1·808	0·842	1·785
0·12	1·616	0·241	2·244	0·157	2·203
0·14	1·357	−0·442	2·639	−0·542	2·572
0·16	1·117	−1·102	2·984	−1·208	2·884
0·18	0·905	−1·710	3·276	−1·812	3·133
0·20	0·725	−2·252	3·516	−2·336	3·320
0·22	0·570	−2·722	3·704	−2·771	3·446
0·24	0·455	−3·117	3·845	−3·115	3·514
0·26	0·357	−3·439	3·942	−3·368	3·528
0·28	0·278	−3·693	4·001	−3·537	3·494
0·30	0·216	−3·884	4·026	−3·625	3·416
0·35	0·113	−4·130	3·968	−3·543	3·058
0·40	0·058	−4·122	3·786	−3·116	2·523
0·45	0·030	−3·944	3·528	−2·447	1·868
0·50	0·015	−3·662	3·228	−1·622	1·137
0·55	0·008	−3·326	2·912	−0·711	0·367
0·60	0·004	−2·969	2·598	+0·231	−0·414
0·7	0·001	−2·276	2·014	2·065	−1·928
0·8		−1·681	1·519	3·673	−3·296
0·9		−1·209	1·123	4·966	−4·460
1·0		−0·852	0·817	5·924	−5·403
1·1		−0·591	0·587	6·571	−6·133
1·2		−0·405	0·418	6·946	−6·669
1·4		−0·186	0·207	7·076	−7·266
1·6		−0·079	0·101	6·674	−7·403
1·8		−0·035	0·048	6·006	−7·251
2·0		−0·015	0·023	5·244	−6·932
2·2		−0·006	0·011	4·488	−6·529
2·4		−0·002	0·005	3·788	−6·089
2·6		−0·001	0·002	3·167	−5·646
2·8			0·001	2·628	−5·213
3·0				2·170	−4·803
3·2				1·785	−4·419
3·4				1·464	−4·061
3·6				1·198	−3·733
3·8				0·979	−3·431
4·0				0·799	−3·154
4·5				0·479	−2·562
5·0				0·286	−2·089
5·5				0·171	−1·710
6·0				0·102	−1·403

Table I—(continued).

r.	(1s).	(2s).	(2p).	(2s).	(3p).
7				0·036	−0·954
8				0·013	−0·654
9				0·004$_5$	−0·451
10				0·001$_5$	−0·313
12					−0·152
14					−0·075
16					−0·037
18					−0·019
20					−0·009
22					−0·004
24					−0·002
26					−0·001

	ϵ	$\int_0^\infty P^2 dr$.		ϵ.	$\int_0^\infty P^2 dr$.
(1s)	209·0	0·5346	(3s)	1·079	67·89
(2s)	18·35	7·252	(3p)	0·114	119·90
(2p)	14·37	7·232			

Table II.—Cl$^-$ Contributions to Z.

Table of $2(2l + 1)[1 - Z_0(nl, nl \mid r)]$.

	(1s)	(2s)	(2p).	(3s).	(3p).
0·000	2·000	2·000	6·000	2·000	6·000
0·005	1·999	2·000	6·000	2·000	6·000
0·010	1·990	1·999	6·000	2·000	6·000
0·015	1·971	1·998	6·000	2·000	6·000
0·020	1·940	1·996	6·000	1·999	6·000
0·03	1·840	1·988	5·999	·999	6·000
0·04	1·700	1·979	5·998	·998	6·000
0·05	1·535	1·969	5·995	997	6·000
0·06	1·358	1·958	5·990	996	5·999
0·07	1·181	1·949	5·982	995	5·999
0·08	1·013	1·942	5·969	994	5·998
0·10	0·717	1·934	5·928	993	5·995
0·12	0·488	1·932	5·859	992	5·991
0·14	0·323	1·931	5·760	1·992	5·985
0·16	0·208	1·927	5·628	1·992	5·978
0·18	0·132	1·916	5·464	1·991	5·969
0·20	0·082	1·894	5·273	1·988	5·959
0·22	0·051	1·859	5·056	1·984	5·948
0·24	0·031	1·812	4·820	1·979	5·935
0·26	0·018	1·753	4·568	1·973	5·923
0·28	0·010	1·682	4·305	1·966	5·910
0·30	0·006	1·603	4·037	1·958	5·899

Table II—(continued).

r.	(1s).	(2s).	(2p)	(3s)	(3p).
0·35	0·002	1·379	3·370	1·939	5·873
0·40		1·142	2·743	1·922	5·853
0·45		0·917	2·187	1·911	5·840
0·50		0·716	1·712	1·905	5·835
0·55		0·548	1·321	1·903	5·833
0·60		0·411	1·005	1·902	5·833
0·7		0·221	0·564	1·898	5·825
0·8		0·114	0·305	1·872	5·791
0·9		0·056	0·161	1·816	5·713
1·0		0·027	0·083	1·728	5·591
1·1		0·013	0·042	1·612	5·423
1·2		0·006	0·021	1·476	5·217
1·4		0·001	0·004	1·182	4·725
1·6			0·001	0·900	4·180
1·8				0·662	3·641
2·0				0·476	3·137
2·2				0·336	2·682
2·4				0·236	2·284
2·6				0·164	1·940
2·8				0·114	1·644
3·0				0·080	1·394
3·2				0·058	1·181
3·4				0·042	1·002
3·6				0·032	0·850
3·8				0·025	0·721
4·0				0·019	0·613
4·5				0·006	0·409
5·0				0·002	0·275
5·5				0·001	0·185
6·0					0·125
7					0·057
8					0·024
9					0·010
10					0·003
11					0·001

§ 5. *Results for* Cu^+.

For Cu^+ the (3d) wave function is very sensitive to the estimated field, the reason being rather different from the case of the (3p) wave function of Cl^-. Here the value of ε is not inconveniently small, but there is a considerable range over which the attractive field is only slightly larger than the centrifugal field (whose potential energy is represented by the term $l(l+1)/r^2$ in the radial wave equation (2)) so that a small change in the attractive field makes a large proportional change in the resultant field and so in the wave function. Analytically, in the wave equation

$$P'' + [2v - \varepsilon - l(l+1)/r^2]\, P = 0.$$

the coefficient of P remains small over a considerable range.

This feature is closely related to the position of Cu in the periodic table. The occurrence of the " transition group " of elements to Sc to Ni in the periodic table is associated with the transition of the (3*d*) wave function from that of potassium, where it is nearly hydrogen-like and is an excited state of the series electron, to the much more compact (3*d*) wave function, forming part of the core, of the copper atom This transition, which is just completed in copper, is due, roughly speaking, to the variation with atomic number of the range over which the attractive force exerted on an electron by the rest of the atom is greater than the centrifugal force on an electron with angular momentum specified by $l = 2$; for potassium the centrifugal force is greatest except at distances well outside the atom ; for copper the atomic field is enough, but not much more than just enough, to hold an electron with $l = 2$ in the core against the centrifugal force.

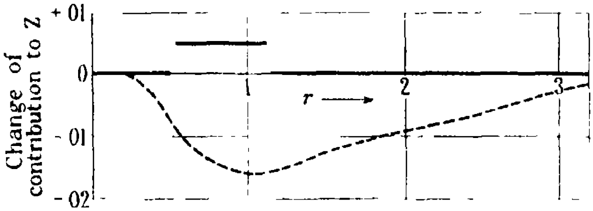


FIG. 1 —To illustrate overstability of $(3d)^{10}$ group. —— Full-line curve shows change of estimated contribution to Z; ---- broken curve shows consequent change of calculated contribution to Z.

This sensitiveness of the (3*d*) wave function, and the comparatively large number of electrons with this radial wave function, makes the $(3d)^{10}$ group seriously over-stable in the sense defined at the end of § 2. Fig. 1 shows quantitatively an example ; the full curve shows a variation made in the estimated contribution to Z from the $(3d)^{10}$ group, and the broken curve the consequent variation in the calculated contribution, which can be seen at once to be considerably greater than the variation in the estimated contribution.

This pronounced overstability of the $(3d)^{10}$ group makes the process of approximation to the self-consistent field very troublesome ; altogether, nine steps of approximation were made in order to obtain the results here given ; most of these steps were mainly concerned with revision of the estimated contribution from the $(3d)^{10}$ group. In the final approximation, for which the results are tabulated, the differences between estimated and calculated contributions to Z from each group are less than 0·01 throughout, the $(3d)$ wave function, and contribution of the $(3d)^{10}$ group to Z, are the best that can be obtained without doing the calculation throughout with an accuracy corresponding to three full decimals in the estimated contributions to Z.

Observed values of ν/R for the X-ray levels are given in the same table as the values of the energy parameter ε of the central field wave functions; although the values of ε cannot be regarded as calculated values of the X-ray levels, as already mentioned, they agree very closely with the observed values for the X-ray levels, a similar close agreement was also found for Rb^+.*

Table III.—Cu^+ Wave Functions.

Table of P/r^{l+1} for small r.

r.	$(1s)$.	$(2s)$	$(2p)$	$(3s)$.	$(3p)$.	$(3d)$.
0·000	100·0	100·0	1000	100	1000	1000
0·005	86·51	86·31	930	86·29	930	953
0·010	74·87	74·12	866	74·04	866	909
0·015	64·80	63·29	807	63·12	806	868
0·020	56·60	53·65	752	53·38	750	828

Table of P.

r.	$(1s)$	$(2s)$.	$(2p)$	$(3s)$	$(3p)$.	$(3d)$.
0·000	0·000	0·000	0·000	0·000	0·000	0·000
0·005	0·433	0·432	0·023	0·431	0·023	0·000
0·010	0·749	0·741	0·087	0·740	0·087	0·001
0·015	0·972	0·948	0·182	0·947	0·181	0·003
0·020	1·122	1·073	0·301	1·068	0·300	0·007
0·025	1·215	1·127	0·438	1·118	0·436	0·012
0·03	1·263	1·125	0·589	1·110	0·585	0·020
0·035	1·276	1·077	0·748	1·055	0·742	0·031
0·04	1·264	0·992	0·912	0·963	0·902	0·044
0·05	1·187	0·742	1·243	0·697	1·223	0·079
0·06	1·071	0·427	1·565	0·365	1·527	0·126
0·07	0·940	0·081	1·864	0·004	1·800	0·184
0·08	0·809	−0·270	2·132	−0·357	2·034	0·253
0·09	0·686	−0·609	2·367	−0·700	2·225	0·332
0·10	0·574	−0·925	2·565	−1·013	2·369	0·420
0·12	0·392	−1·462	2·855	−1·513	2·518	0·621
0·14	0·261	−1·858	3·013	−1·830	2·495	0·847
0·16	0·170	−2·119	3·063	−1·968	2·326	1·091
0·18	0·110	−2·263	3·027	−1·946	2·038	1·346
0·20	0·070	−2·311	2·925	−1·793	1·659	1·605

---

* Hartree, *loc. cit.*

Table III—(continued).

r.	(1*s*).	(2*s*).	(2*p*).	(3*s*).	(3*p*).	(3*d*).
0·22	0·044	−2·286	2·779	−1·536	1·216	1·863
0·24	0·027	−2·206	2·602	−1·202	0·731	2·117
0·26	0·017	−2·089	2·409	−0·817	0·223	2·362
0·28	0·010	−1·948	2·208	−0·400	−0·290	2·595
0·30	0·006	−1·795	2·007	+0·030	−0·797	2·815
0·35	0·002	−1·400	1·534	+1·079	−1·974	3·296
0·40		−1·044	1·136	1·988	−2·953	3·671
0·45		−0·754	0·821	2·694	−3·701	3·944
0·50		−0·533	0·583	3·190	−4·222	4·126
0·55		−0·369	0·408	3·494	−4·545	4·231
0·60		−0·253	0·282	3·637	−4·702	4·276
0·7		−0·115	0·132	3·578	−4·654	4·236
0·8		−0·051	0·061	3·247	−4·314	4·089
0·9		−0·022	0·028	2·804	−3·839	3·888
1·0		−0·009	0·012	2·349	−3·328	3·666
1·1		−0·004	0·005$_5$	1·925	−2·834	3·439
1·2		−0·002	0·002$_5$	1·553	−2·382	3·217
1·3		−0·001	0·001	1·238	−1·984	3·004
1·4				0·978	−1·640	2·803
1·6				0·599	−1·103	2·437
1·8				0·359	−0·731	2·117
2·0				0·213	−0·479	1·838
2·2				0·125	−0·311	1·594
2·4				0·073	−0·201	1·381
2·6				0·042	−0·129	1·196
2·8				0·024	−0·083	1·034
3·0				0·014	−0·053	0·892
3·2				0·008	−0·034	0·769
3·4				0·004	−0·021	0·662
3·6				0·002	−0·014	0·568
3·8				0·001	−0·009	0·487
4·0					−0·005	0·417
4·5					−0·001	0·281
5·0						0·188
5·5						0·124
6·0						0·081
7						0·034
8						0·014
9						0·006
10						0·002$_5$
11						0·001
12						

	ϵ.	ν/R obs. X-ray term.	$\int_0^\infty P^2\,dr$		ϵ	ν/R obs X-ray term	$\int_0^\infty P^2\,dr$
(1*s*)	658·0	661·6	0·10589	(3*s*)	8·968	8·9	7·905
(2*s*)	78·45	81·0	1·1568	(3*p*)	6·078	5·7	14·913
(2*p*)	69·86	68·9	2·0621	(3*d*)	1·195	0·4	21·296

Table IV.—Cu^+ Contributions to Z.

Table of $2(2l+1)[1-Z_0(nl, nl|r)]$.

r.	$(1s)^2$.	$(2s)^2$.	$(2p)^6$.	$(3s)^2$.	$(3p)^6$.	$(3d)^{10}$.
0·000	2·000	2·000	6·000	2·000	6·000	10·000
0·005	1·993	1·999	6·000	2·000	6·000	10·000
0·010	1·958	1·996	6·000	1·999	6·000	10·000
0·015	1·887	1·990	6·000	1·998	6·000	10·000
0·020	1·782	1·981	5·999	1·997	6·000	10·000
0·025	1·652	1·970	5·997	1·996	6·000	10·000
0·030	1·506	1·959	5·993	1·994	5·999	10·000
0·035	1·354	1·949	5·987	1·992	5·998	10·000
0·040	1·201	1·940	5·977	1·991	5·997	10·000
0·05	0·915	1·926	5·942	1·990	5·992	10·000
0·06	0·674	1·920	5·884	1·989	5·984	10·000
0·07	0·482	1·919	5·798	1·989	5·973	10·000
0·08	0·338	1·918	5·681	1·989	5·958	10·000
0·09	0·232	1·915	5·533	1·988	5·940	10·000
0·10	0·157	1·904	5·355	1·986	5·919	9·999
0·12	0·069	1·854	4·923	1·977	5·870	9·996
0·14	0·029	1·757	4·418	1·963	5·819	9·991
0·16	0·012	1·618	3·878	1·945	5·772	9·982
0·18		1·451	3·336	1·925	5·733	9·968
0·20		1·268	2·819	1·907	5·705	9·947
0·22		1·085	2·345	1·893	5·688	9·919
0·24		0·910	1·922	1·883	5·680	9·882
0·26		0·750	1·558	1·878	5·678	9·835
0·28		0·609	1·248	1·876	5·678	9·776
0·30		0·487	0·990	1·876	5·675	9·708
0·35		0·266	0·534	1·871	5·633	9·487
0·40		0·137	0·277	1·839	5·507	9·200
0·45		0·068	0·138	1·768	5·282	8·858
0·50		0·032	0·067	1·657	4·963	8·474
0·55		0·015	0·032	1·514	4·573	8·062
0·60		0·007	0·015	1·353	4·141	7·637
0·7		0·001	0·002	1·018	3·249	6·781
0·8				0·721	2·433	5·965
0·9				0·488	1·762	5·217
1·0				0·320	1·243	4·547
1·1				0·204	0·861	3·953
1·2				0·128	0·587	3·433
1·3				0·079	0·396	2·980
1·4				0·047	0·265	2·584
1·6				0·016	0·115	1·939
1·8				0·005	0·048	1·454
2·0				0·001	0·019	1·087
2·2					0·006	0·811
2·4					0·002	0·604
2·6						0·448
2·8						0·330
3·0						0·243
3·2						0·179
3·4						0·130
3·6						0·095
3·8						0·068
4·0						0·050

Table IV—(continued).

r.	$(1s)^2$.	$(2s)^2$.	$(2p)^6$.	$(3s)^2$.	$(3p)^6$.	$(3d)^{10}$.
4·5						0·022
5·0						0·008
5·5						0·003
6·0						0·001

It is interesting to compare the distribution of charge for Cu^+, with an outer shell of 18 electrons, with that for inert-gas-like ions which have outer shells of 8 electrons; and for this purpose the total charge density for Cu^+ and contributions to it from the groups of the outer shell† are shown as functions of r in fig. 2, and corresponding diagrams for K^+ and Rb^+ are shown in figs. 3 and 4 respectively.‡

The Cu^+ ion is more compact than K^+ or Rb^+; although the outer shell of Cu^+ consists of 18 electrons while those of K^+ and Rb^+ consist of 8, the charge density in Cu^+ is smaller than in K^+ or Rb^+ outside a radius of about 1·2 atomic units; and it is necessary to pass right inside the outer shell of K^+ or Rb^+ to the next inner shell before coming to a region of charge density as high as that in the outermost shell of Cu^+. This agrees with the much smaller observed packing radius for a Cu^+ or Cu^{++} ion than of a K^+ or Rb^+ ion in crystals (the effect of a second ionization of the copper atom would be to contract the "tail" of the $(3d)$ wave function for large r, but it would not greatly alter the distribution).

It should be pointed out here that the total volume density, which is $1/4\pi r^2$ times the radial density, increases steadily as r decreases, and does not show the maxima and minima shown by the radial density curve.

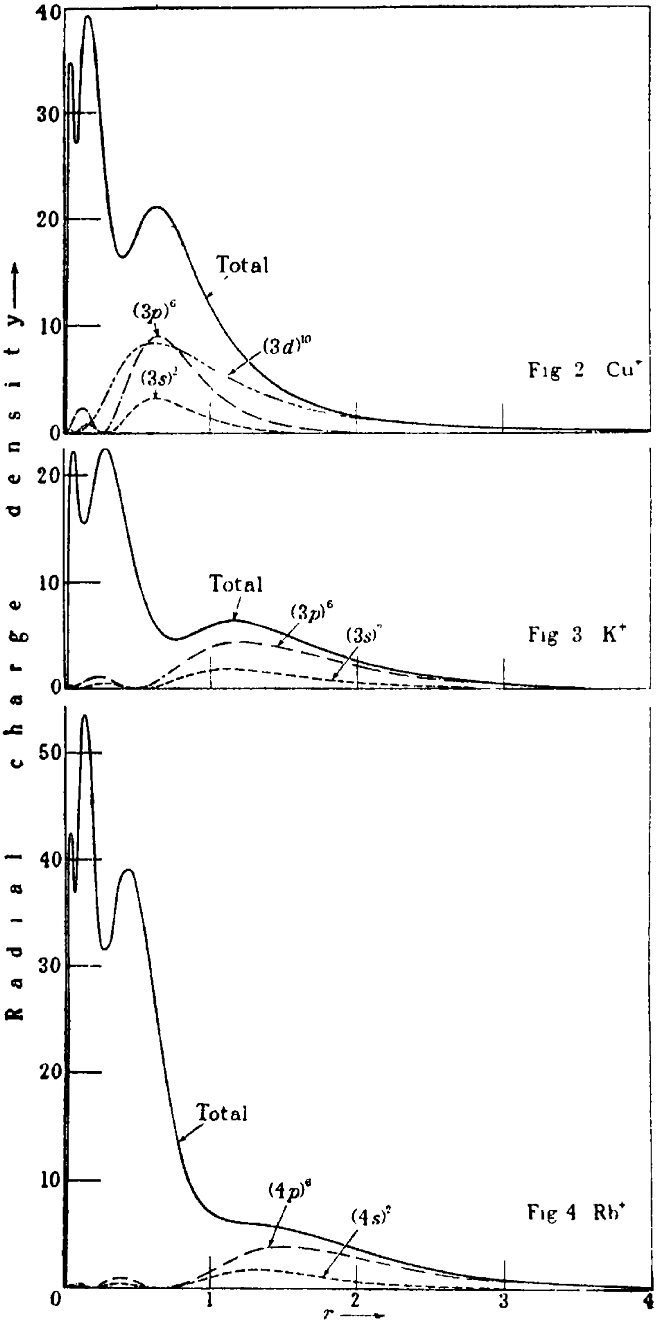
Another interesting comparison between Cu and the alkali metal atoms is that of the series electron wave function and its relation to the interatomic distance in metals. This comparison will be given here although it depends on some results not given in this paper. Slater's theory of cohesion in metals§ suggests that the interatomic distance in the crystal of a metallic element

---

† To avoid crowding the diagram, the contributions from the individual groups of the inner shell are not shown; the point of interest in the comparison is the structure of the outer shell.

‡ These diagrams are based on results of old work, but the alteration in the results due to revision based on the standard calculations is not likely to be appreciable on the scale of these diagrams.

§ Slater, 'Phys. Rev.,' vol. 35, p. 509 (1930).



Figs. 2, 3, 4.—Charge distributions in Cu+, K+, Rb+. Radial charge density in electrons per atomic unit shown as function of radius in atomic units. The full-line curve shows total charge density for the whole atom, the broken curves show the contributions to the total charge density from the groups of the outer shell.

may be given approximately by the condition that the unperturbed series electron wave functions of the two atoms overlap so that their maxima nearly coincide.

If P is the radial function for the series electron, then the wave function itself is P/r (the normal state of the series electron being an *s* state); in Table V the approximate radii of the main maxima of P^2 and of $(P/r)^2$ for the normal series electron wave function are given; and also the observed atomic packing radius (defined as half the interatomic distance) and its ratio to the radii of the maxima of the calculated wave functions.

Table V.

Atom.	Radius of max P^2, r_1.	Radius of max $(P/r)^2$ r_2	Crystal structure of element.	Packing radius* r_0.	$\frac{r_0}{r_1}$	$\frac{r_0}{r_2}$
K	2·66	2·01	Body centred cubic	2·309	0·87	1·15
Rb	2·84	2·18	,,	2·43	0·86	$1{\cdot}11_5$
Cs	3·21	2·53	,,	2·62	0·81	1·04
Cu	1·51	1·04	Face centred cubic	1·275	$0{\cdot}84_5$	1·23

(All lengths in A.)

* Taken from Neuberger, 'Z. Krystallog.,' vol. 80, p. 103 (1931).

The results show a relation between series electron wave functions and interatomic distance of the kind suggested by Slater; the atomic radius in all cases lies between the radii at which P and P/r have their maxima, and, except for Cs, about half way between, and the variation of either ratio is considerably less than the variation of the packing radius.

§ 6. *Summary.*

Approximate wave functions for a number of atoms have been calculated by the method of the self-consistent field, but of the available results few have been published in detail. These results for a number of atoms are being revised, and the approximation improved, with a view to publication so as to make them more generally accessible; this paper presents the results for Cl^- and Cu^+. The results for copper are compared with those for the other alkali metals and characteristic differences noted; and the relation between the interatomic distance in metals and the wave function of the series electron is pointed out.

Laboratory Determinations of the Magnetic Properties of Certain Igneous Rocks.

By A. F. HALLIMOND, M.A., Sc.D., with contributions by Professor E. F. HERROUN.

(Communicated by Sir Frank Smith, Sec. R.S —Received January 24, 1933)

Although the methods of magnetic survey in the field are now well established, there have been relatively few direct determinations of the magnetic properties of the rocks concerned. Conclusive evidence has now been obtained that some of the natural rock-masses are permanently magnetized ; this effect is superimposed upon the magnetization induced by the earth's field, and in certain rocks it appears that even a small permanent magnetization may seriously modify the magnetic profiles. Experiments were therefore undertaken with the general object of linking up the evidence provided by field surveys with that obtainable in the laboratory.

The material was chosen with a view to providing an experimental basis for the interpretation of the magnetic surveys that have been completed by the Geological Survey during the past few years. The collection and preparation of the specimens were undertaken by the Survey, and the laboratory measurements have been carried out by Professor E. F. Herroun, to whom the author has been greatly indebted for criticism and advice throughout this work.

Violent magnetic disturbances are caused by electric railways and similar circuits, and on this account work in very low fields has become difficult even at many research institutions; fortunately it was possible to obtain the necessary quiet conditions in Professor Herroun's laboratory at Reigate. The present report contains an account of the apparatus used and of the state of magnetization observed in a series of cubes of natural rock. The question whether permanent magnetization of the order observed can affect the magnetic profiles is discussed with special reference to the magnetic survey of the Lornty Dyke.

1. *The Magnetometer.*

Measurements were made by reading the deflection produced on placing cubes of rock, of about 3·3 cm. side, at a distance of a few centimetres from a freely suspended magnetized needle.

The cube must necessarily be of fair size in order to obtain a measureable magnetic moment; owing to its low intensity of magnetization it must be placed rather near the needle of the magnetometer, in which position its pole-face subtends a considerable solid angle at the distance of the needle. This face can hardly be treated as a magnetic shell and its surface may not be uniformly magnetized. The latter condition is due partly to the fact that the surfaces of the cube, as cut, are not normal to the axis of magnetization, and partly to the distribution of the active constituent (magnetite) not being always uniform throughout its volume.

It was therefore considered best, in order to minimize these errors, to place the cube as far from the needle as possible, consistent with reliable deflections being obtained with the mirror and scale; and further to increase the sensitivity by the use of a small compensating magnet placed on the side of the needle remote from the specimen so as to weaken the restoring force of the earth's field at the needle.

Accordingly the cubes were carefully placed centrally on a horizontal brass turn-table, the centre of which was 7·0 cm. from the suspended needle; on the other side of the needle on the E -W line the small compensating magnet was placed with its axis in the magnetic meridian and by a very small movement in the N—S direction the zero of the scale reading could be made to coincide with the original zero in the earth's field.

The value of the reduced field at the needle was compared with the full field due to H both by finding the squares of the periodic times of oscillation of the needle and also by the deflection method with and without the compensating magnet. In this way the reduced field at the needle was found to be from 0·067 to 0·07 C.G.S. unit under the various conditions in which the instrument was employed.

It is evidently undesirable to use much smaller cubes than those described and to place these very close to the needle, since the latter when deflected might induce polarity in the specimen and give misleading, high, results.

But the uncertainty of the distance between the effective poles with a cube of 3·3 cm. in the side is a drawback; moreover, the usual method of eliminating all but the higher terms of $l/2$ by taking two distances, say in the ratio of 3 · 4, is not applicable to the cubes as received, for with any material increase in distance the deflection falls to such an extent as to render its reading unreliable. With an artificially magnetized cube, however, good deflections are obtainable at 15 or 20 cm. from the needle, but this distance is too great for the feebly magnetized cube in its natural state.

By the use of small bars of hard steel, weakly magnetized, and a small cut bar of magnetite of known magnetic moment, constants for this instrument were obtained for the 7 cm. distance. The cubes were all substantially of the same dimensions, they were placed centrally on the turn-table and readings were obtained on rotating the cube through 180°. It is believed, therefore, that the results are strictly comparable with each other and that they represent the true values to within 3 or 4 per cent.

A magnetizing solenoid and compensating coil were also used with this magnetometer; when a cube of Cleveland Dyke rock was put inside the solenoid and a current giving a calculated value of H within the solenoid of 1·186 C.G.S. was switched on and reversed, deflections were obtained giving $M = 0{\cdot}103$, $J = 0{\cdot}00299$ and $K_v = 0{\cdot}00252$, a result in very good agreement with values found by instruments of the type of the Curie balance for somewhat higher fields.

It is evident that, with such a low K_v value, the "end effects," even with cubes, must be small.

II. *Permanent Magnetization of Rocks* in situ.

The Effects of Lightning.—Ground struck by lightning is penetrated to a depth of many feet, the path of the discharge being indicated in extreme cases by tubes of fused rock termed "fulgurites." It may be anticipated, therefore, that pieces of rock collected from an old land surface will have been magnetized at some period in their history. Loewinson-Lessing* has described a very marked instance of this effect: andesite from the summit of the mountain Petit Ararat has been penetrated by fulgurites and is intensely magnetized, to a degree far greater than that produced by cooling in the earth's field.

Lightning discharges are less frequent in open country than on mountain summits, but there can be little doubt that rocks like that of the Swynnerton Dyke must have occasionally undergone magnetization from this cause, for they have been exposed at the surface of the ground for a very long period of time.

Magnetization by lightning is virtually instantaneous; indeed Pockels† has made use of the effect to record the current in a conductor carrying a lightning discharge; small bars of basalt were placed in the vicinity of the conductor, and their magnetization was subsequently determined.

---

* 'Bull. Acad. Sci. U.S.S.R.,' p. 875 (1927); 'C. R. Acad. Sci. Paris,' vol. 180, p. 242 (1925).

† 'Ann. Chem. Phys.,' vol. 63, p. 195 (1897), and 'Phys. Z.,' vol. 2, p. 334 (1901).

Lightning strokes differ in effect from the earth's field in that the resulting magnetization is not uniform in strength or direction. Not only does the discharge ramify in the ground, but each line of current exerts a circular field, so that pieces of rock on either side of the discharge are magnetized in opposite directions; the total external effect may therefore be relatively small. To produce an appreciable effect at a height of 4 feet a magnetized area exceeding several inches in diameter would be necessary. Tests by Koenigsberger* showed that in most cases the irregular polarity when present varies from point to point within a few inches distance; it is therefore likely that magnetization of this kind is generally too irregularly distributed to affect the magnetic profiles. This would explain the consistency of the results obtained in field surveys, *e.g.*, over the outcrop of the Swynnerton Dykes, and it will apply with even greater force when the rock mass is at a greater distance from the instrument. Lightning strokes, except under abnormally severe conditions, and at a direct outcrop, can probably be ignored in considering the results of a magnetic survey.† In laboratory determinations, on the other hand, where small samples are used, practically all measurements on outcrop material will be subject to large errors from this cause; it therefore seems advisable to use only specimens taken from a considerable depth, say 50 feet.

For the first test of the magnetic condition of a rock taken at depth, specimens were collected from the Cleveland Dyke.‡ This rock is mined at Goathland, N. Yorkshire, as a roadstone; the drift is at a depth of about 100 feet below the open moorland, where the rock outcrop will be subject to considerable risk of lightning stokes. Five rough specimens, each about 4 inches across, selected from the mined material, were tested in front of the magnetometer and gave the following deflections: 25, 20, 15, 10, 25 scale divisions.

The polarity appeared to be somewhat irregular in distribution, but this effect was in no specimen large and it seems clear that the rock at this depth possesses a weak but fairly uniform magnetization. A cube with 3·3 cm. side was next cut from a representative piece of the above set, and was found to be magnetized with the intensity 43×10^{-5}. The Cleveland rock at depth is thus only weakly magnetized. If the susceptibility is taken as 250×10^{-5}, the induced magnetization (kV), due to the earth's field V, will be about 125×10^{-5}, a value nearly three times the permanent magnetization.

---

* 'Beitr. Geophys.,' vol. 35, p. 204 (1932), containing reference to earlier papers.

† These remarks apply to moderate anomalies due to ordinary igneous rocks—hardly to magnetite deposits.

‡ Slide E 16778.

In order to ascertain the direction in which the dyke is magnetized, it would be necessary to take orientated samples from the rock *in situ*; arrangements were accordingly made to do this for the material collected from the Blairgowrie Dyke and from Leicestershire. The last-named rocks proved, however, to be only very weakly magnetic, for reasons to be mentioned below.

Parallel determinations on the Leicestershire, Blairgowrie and other rocks are given in Table I.

Table I.—Permanent Magnetization of Rock Cubes as cut. Intensity of Magnetization J along the cube axes *a*, *b*, *c*.

Specimen.	*a*.	*b*.	*c*.	Susceptibility $\times 10^{-5}$.
Roadstone No. 970, Seacliff Quarry, N Berwick (depth not known)	121	26	69	(*q*) 363
Cleveland Dyke, Goathland, N. Yorks.	22	37	0	(*r*) 253
Tholeite, Cat's Craig Quarry, Blairgowrie	−8	0	25	(*p*) 164, (*q*) 226
Do , 2nd sample	similar values to preceding.			
Greenstone Dyke, Mt. Sorrel, Leics.	0	0	0	(*q*) 36·5
Granophyre, Bradgate Hill Quarry, Groby, Leics (No. 18)	0	0	0	(*r*) 5·4
Granophyre, "New Quarry," Groby	0	0	0	(*q*) 13·3

NOTE.—The susceptibilities are in 10^{-5} C G.S. They were determined as follows: (*p*) on a cut cube, H = 5·0, (*q*) on powder, H = 88; (*q'*) on powder, H = 130; (*r*) on a cut bar, H high.

*Orientation**.—The last five samples were orientated and here the axis *c* was vertical, *b* pointing E—W and *a* N—S (magnetic). The sign given is positive when the upper, the eastern, or the southern end of the sample has "south" polarity.

The Seacliff specimen was selected for a preliminary test, from the collection of roadstones; it may have been at any depth in the quarry. The others were specially taken at depths of 30 feet or more below the surface of the ground.

The values for the Leicestershire rocks are low. This agrees with the weak effects observed in a magnetic survey (not published) over a small area of the Groby granophyre. Equally low laboratory results were obtained for the Mount Sorrel granite; this rock is extensively reddened, and only the grey variety was susceptible.† From the very definite results obtained in a magnetic survey it may be assumed that the granite in depth is susceptible, but in view of the weak and variable laboratory results it has not been possible to get decisive evidence for the present purpose from any of the Leicestershire rocks. The oxidation and loss of magnetic properties may be associated with the fact that this region is an old desert land-surface, in which weathering may have operated to considerable depths.

---

* The pieces collected were loose but still in position in the quarry. The orientated piece was marked with an arrow to indicate the magnetic north, and a line to show the horizontal plane; it was then detached from the quarry face. At the same time six pieces of the rock were collected within about a yard of the marked piece, to check the uniformity of the magnetization.

† *Cf.* measurements by Wilson, 'Phil. Trans.,' A, vol. 219, p. 84 (1919).

III. *Effect of Permanent Magnetization on the Magnetic Profiles.*

The Lornty Dyke.

Magnetic profiles obtained in the survey of the Swynnerton Dyke were symmetrical and gave no indication of any important degree of cross-magnetization. At other localities, however, profiles have been recorded indicating the presence of a considerable effect which might be due to permanent magnetization of the dyke rock. One outstanding case is that recorded by Schulze* for dykes in Saxony. Here the curves, instead of a simple maximum above the magnetic rock, have also a negative portion, so that their general form is that shown in fig. 3 ; the south side of the dyke had apparently a magnetic " south " polarity.

The first example of this kind to be recorded in Great Britain was observed by Drs. McLintock and Phemister in their survey of the Lornty Dyke, near Blairgowrie, in Perthshire.† The Lornty curves suggest permanent magnetization, but the direction is the opposite of that which would be induced by the earth's present field. It was with considerable interest, therefore, that specimens of this rock were tested by the present methods in order to ascertain directly the intensity and direction of the permanent magnetization in the cut cubes. It will be shown that the values observed, though at first sight small, are in the required direction and are of the order required to account for the very well-marked peculiarities of the Lornty profiles.

Six rough specimens and two orientated pieces were collected from Cat's Craig Quarry, the only quarry from which fresh rock was obtainable The depth was about 30 feet. One of the rough pieces proved to be rather strongly magnetic (deflection = 150 divisions) and probably represents a local variation in the dyke ; a microsection (E 29624) shows that the rock is greener and more completely crystalline ; the other pieces (E 29622/3) show confused interstitial patches (mesostasis) containing fine-grained magnetite with a few red flakes of ferric oxide. Otherwise the rocks are very similar. The remaining pieces were moderately magnetic, yielding the deflections 14, 15, 20, 13, 14, so that the magnetization appears, as in the Cleveland Dyke, to be fairly uniform. Two cubes were cut from the special orientated pieces, with axes pointing respectively magnetic north, west and vertical. These when tested on the magnetometer yielded the following values :—

---

* 'Z. Geophys.,' vol. 6, p. 141 (1930).

† 'Summary of Progress for 1930,' Part III (Mem. Geol. Survey), pp. 24-29 (1931).

	Deflection scale divisions	Intensity of magnetization 10^{-5} C.G.S
Axis *c* (vertical)	3	25
a (N—S)	−1 nearly	−8
b (E—W)	0	0

The results for cube 2 were exactly similar.

These values are small but are quite consistent and decisive.

The resultant intensity of permanent magnetization is therefore $J = 26 \times 10^{-5}$ C.G.S.

The direction is approximately in the present magnetic meridian, but inclined about 15° northwards from the vertical with the south pole on the north side of the dyke. It is thus inclined at about 35° to the earth's present field. The intensity of magnetization on the walls of the dyke will be 20, and on the horizontal top of the dyke, 25×10^{-5} C.G.S.

When placed in a solenoid with approximate field-strength 5·0 C.G.S. the susceptibility of one of the cubes was found to be 164×10^{-5} C.G.S.

Comparison of Observed with Calculated Profiles for the Lornty Dyke.—Vertical force profiles were published in 1930.* Three of these are sufficiently complete to serve for the present comparison, and these have been reproduced in fig. 3 ; they were obtained at places about half a mile apart along the dyke. No. II resembles the simple profiles met with at Swynnerton,† and can be used in the same way to give a rough estimate of the depth. It was pointed out, in discussing the Swynnerton profiles, that such an estimate could only be made in the absence of disturbances due to magnetization of the sides of the dyke. Such disturbances are evidenced in a profile by lack of symmetry between the two sides, and attention was drawn to the presence of this effect to a small extent in one of the Swynnerton traverses. The remaining two traverses now under discussion, Nos. I and III at Lornty, offer an outstanding example of the same asymmetry, which, it will be shown, can be accounted for with considerable accuracy on the assumption that the dyke is permanently magnetized. The susceptibility measured on the cut cube of Lornty rock agrees with that required by all three profiles, while the small degree of permanent magnetization measured on the cubes is the same in direction though somewhat smaller in amount than that required by the profiles for the dyke as a whole.

---

* McLintock and Phemister, *loc. cit.*

† Hallimond, "Summary of Progress" (' Mem. Geol. Surv.'), Part III, p. 44 (1929).

Profiles of the unsymmetrical type (like I and III) were observed by Schulze* over E—W basalt dykes in Saxony, and theoretical profiles were calculated by way of the gravitational potential. For the Lornty Dyke the method of calculation follows that outlined by the present author for the Swynnerton Dyke. The geological structure is represented by a model constructed of plane surfaces separating regions of uniformly magnetized material; each surface has a uniform polarity, and the effect of each surface is calculated for the point of observation. Thus the magnetic anomaly is obtained as the sum of a number of terms each representing one of the surfaces. This method has the advantage of assigning a direct physical meaning to each term, and it is correspondingly easy to estimate the effect upon the profile when the model is altered either in shape or in material.

Vertical force profiles will be calculated in this way for the three Lornty traverses.

General Formulæ for the Profiles above an Inclined Dyke.—The dyke is regarded as a parallel-sided layer of material with susceptibility k, dipping at an angle ϕ to the horizon and extending indefinitely downwards and to the sides. The top is a horizontal infinite strip, of width w.

(1) Magnetic profiles due to the top of a wide dyke: horizontal and vertical forces at a point P above a horizontal infinite strip of width w, intensity of magnetization i. Notation as in fig. 1; AB is the trace of the strip.

From the calculation already made for the Swynnerton Dyke,† the forces at a point P due to the elementary infinite strip whose trace is dx, are

$$v = 2i \cos \theta \, dx/l \qquad h = 2i \sin \theta \, dx/l.$$

Integrating across the strip AB,

$$v = \int_{\theta_1}^{\theta_2} 2i \, d\theta = 2i (\theta_2 - \theta_1),$$

$$h = \int_{\theta_1}^{\theta_2} 2i \tan \theta d\theta = 2i (\log \sec \theta_2 - \log \sec \theta_1).$$

---

* *Loc. cit.* For calculation see H. Haalck, "Die magnetischen Verfahren" (Berlin, 1927); Slichter, 'Amer. Inst. Min. & Met. Eng., Tech. Pub.,' No. 120, p. 17 (1928) points out that one of Haalck's formulæ is erroneous.

† "Summary of Progress for 1929," Part III, 'Mem. Geol. Surv.,' p. 51 (1930).

These formulæ give the vertical and horizontal force profiles for traverses at right angles to the dyke. The vertical force expression is very simple, being proportional to the angle subtended by the horizontal strip. It is very widely useful, both for the top of a dyke and for the concealed boundary of a magnetic formation, when the face of contact is not magnetized.*

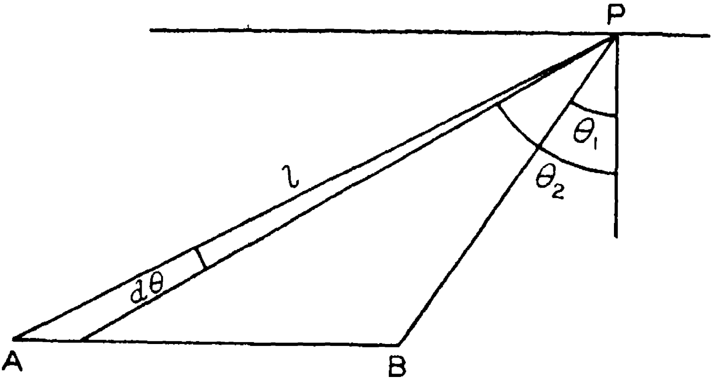


FIG. 1.—Magnetic forces due to a horizontal strip.

(2) The sides of the dyke.

(*a*) Vertical and horizontal forces at a point P above an inclined plane strip. Notation as in fig. 2. AB is the trace of an infinite strip of width w inclined at an angle ϕ, but with its edges horizontal.

The forces at P perpendicular and parallel to the plane of the strip have already been obtained in (1). They are

$$\text{Force perpendicular to AB} = 2i(\theta_2 - \theta_1)$$

$$\text{Force parallel to AB} = 2i \log \frac{\sec(\theta_2 + \frac{1}{2}\pi - \phi)}{\sec(\theta_1 + \frac{1}{2}\pi - \phi)} = 2i \log r_2/r_1.$$

The vertical and horizontal forces at P are therefore

$$v = 2i \log r_2/r_1 \sin\phi - 2i(\theta_2 - \theta_1)\cos\phi$$
$$h = 2i \log r_2/r_1 \cos\phi - 2i(\theta_2 - \theta_1)\sin\phi.$$

(*b*) Magnetic profile due to the two sides of an inclined dyke. The horizontal and vertical forces calculated in the preceding paragraph for one side of the dyke alone, become infinite when the width of the strip is indefinitely increased ($r_2 = \infty$). When there are two sides (with opposite polarity) the distant portions of the strips neutralize one another, and the magnetic force at P is still finite. Using the notation r'_2, r'_1, etc., for one strip and r''_2, r''_1, etc., for the other strip, when $r_2 = \infty$; $r'_2/r''_2 = 1$, $\theta'_2 = \theta''_2 = \phi$, and the forces are given by

---

* In the case of an isolated boundary the width of the "dyke" is infinite and the formula becomes $v = 2i(\pi/2 - \theta_1)$; for a point above the magnetic rock distant from the boundary, $v = 2\pi i$; for a point directly above the boundary $v = \pi i$; at a distance d from the boundary, $\theta = 45° = \pi/4$ and $v = \frac{1}{2}\pi i$. Thus d can be obtained directly from the profile.

$$v = v' + v'' = 2i \sin \phi \log r''_1/r'_1 - 2i \cos \phi(\theta''_1 - \theta'_1)$$
$$h = h' + h'' = 2i \cos \phi \log r''_1/r'_1 - 2i \sin \phi(\theta''_1 - \theta'_1).$$

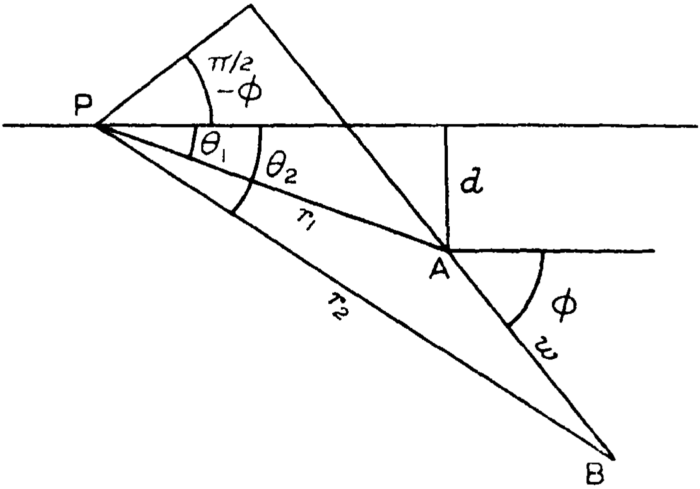


FIG. 2.—Magnetic forces due to an inclined strip.

(3) Magnetic profiles for the complete dyke. The magnetic profile for a traverse across the dyke is given by the sum of the forces due to the top and sides. These in general have different values for the intensity i, which is in each case the sum of the magnetization induced by the earth's field and the component of the permanent magnetization. It is convenient, therefore, to calculate them separately and plot the sum of the values.

Calculated Profiles for the Lornty Dyke.—Laboratory measurement of the susceptibility has provided the value 165×10^{-5}. The magnetization induced by the earth's field in the top of the dyke should therefore be $kV = 0{\cdot}46 \times 165 \times 10^{-5} = 76 \times 10^{-5}$. To this must be added the component of the permanent magnetization, which is roughly 25; so that the total magnetization of the top should be about $i = 100 \times 10^{-5}$. The sides are believed to dip* nearly parallel with the earth's field (at about 70°) so they will have no induced magnetization, but the laboratory values indicate a permanent magnetization of about 20 units, the north side being south polar.

---

* This dip is well proved by the quarry sections.

The width of the dyke is known to be about 30 feet; profiles were accordingly constructed with width 30, dip 65° and depth (*a*) 6 feet, (*b*) 30 feet. The values proved to be of the right order, but one alteration had to be made before the curves could be brought into numerical agreement with the profiles measured in the field survey. It was found necessary to regard the sides of the dyke as magnetized in the direction already indicated, but with a greater intensity, of about 100×10^{-5}. The theoretical curves for these values are shown superposed upon the observed profiles in fig. 3 The agreement is fairly good both in form and in magnitudes, the only serious deviation being the rather high observed values to the north of the outcrop. In traverse I the curves would indicate that the dyke is a little wider than 30 feet and the cover a little thicker than 6 feet. Traverse II is not at right angles to the dyke and the peak is consequently rather widened; the depth indicated is about 30 feet and the width 30 feet. Traverse III is over a direct outcrop and is in very close agreement, the dyke being apparently just under 30 feet wide.

Nearly all dykes contain patches of more basic material, usually near the sides. These may be erratic in their distribution, and the principal difference in the curves (which suggests more magnetic material in the deeper part of the dyke) might be explained either by basic patches or by incipient weathering of the top part of the dyke.

Self-magnetization.—The preceding calculation ignores disturbances due to the magnetization of the dyke in its own field. The corners of the dyke are subject to a field due to the polarity distributed over the top and sides. The direction and amount of this field outside the dyke have been calculated above (assuming uniform magnetization); at a distance of 6 feet above the corner the force is of the order of 0·00500 C.G.S. But the permeability is only 0·0016, so that the magnetizing effect of the rest of the dyke on a piece of dyke material at 6 feet away would be negligible. Any edge effect might be expected to become noticeable only within a foot or less of the extreme corner of the dyke. It is difficult to suppose that this will account for the apparently enhanced value of the permanent magnetization as indicated in the profiles, though some allowance may be required. The disturbance would apparently take the form of roughly diagonal magnetization outcropping in narrow strips along the corners. The effects of top and sides reinforce on the northern corner and neutralize on the southern, so that it seems likely that the general effect would be to accentuate the positive peak of the profile.

Summary.

An account is given of the magnetic properties of igneous rock specimens collected by H.M. Geological Survey. The rocks were taken from localities at which field surveys have been made with the magnetic vertical force variometer during the past few years. A group of several specimens was taken at each place in order to ensure uniformity, and from one of these a 3·3 cm. cube was

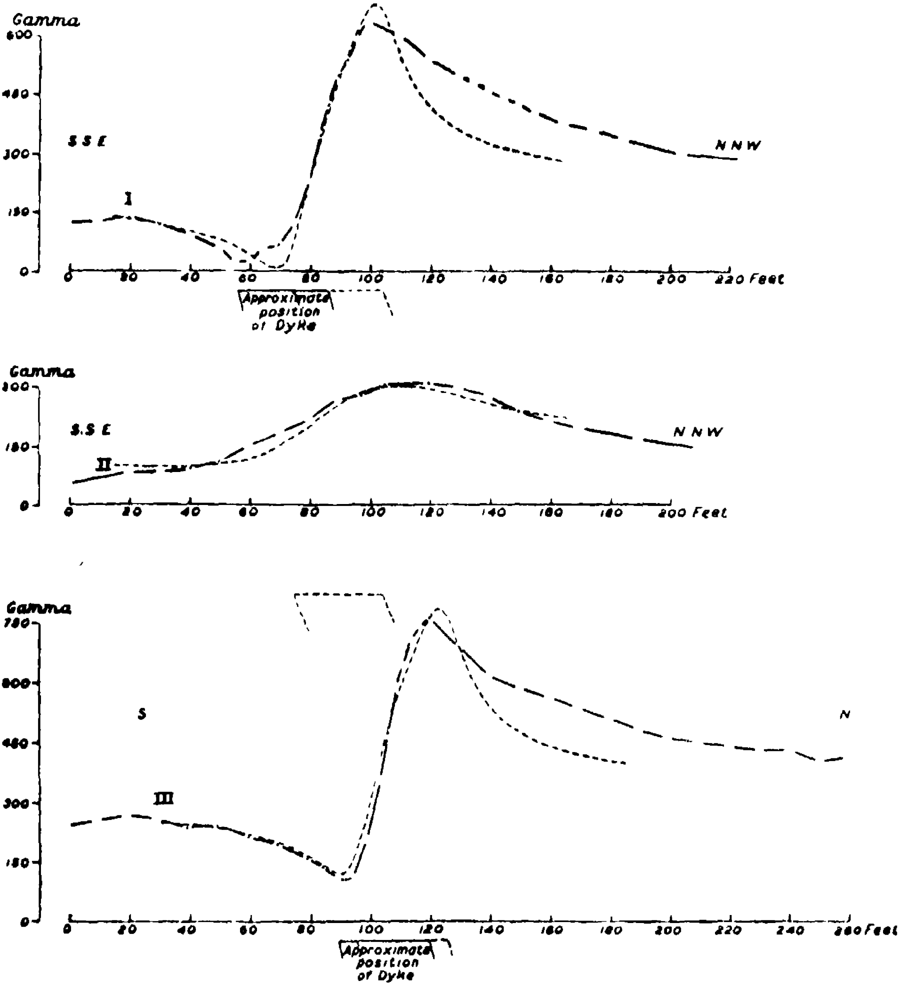


FIG. 3.—Observed and calculated vertical force profiles across the Lornty Dyke. The "calculated" profiles and corresponding position of the dykes are shown by broken lines. Points on the "observed" curves represent magnetometer readings.

cut for magnetic determination in the laboratory. The measurements, which were made by Professor Herroun, were obtained by placing the cube near a freely suspended magnetized needle, in a partially neutralized field. The instrument was standardized by means of bars of known moment and by the time of swing. Susceptibility was measured by surrounding the cube with a coil.

Specimens from Leicestershire (Charnwood Rocks) gave low values, but a measureable degree of permanent magnetization and susceptibility was found for the Cleveland Dyke and for the dyke at Lornty in Perthshire. A method is developed for calculating the magnetic profile from the geological model assuming uniformity of magnetization. Calculated and observed curves for the Lornty profiles show a fairly good agreement; they indicate that the north side of the dyke is south-polar, so that the magnetization appears to be in the opposite sense to that recorded for certain dykes in Germany, and does not coincide with that which would be produced by cooling in the earth's present field.

New Method of Deriving Stresses Graphically from Photo-Elastic Observations.

By H. Neuber,* Munich.

(Communicated by E. G. Coker, F.R.S.—Received February 4, 1933.)

In transparent bodies under plane stress the difference and directions of the two principal stresses are obtained by observations with polarized light. To separate the stresses either additional measurements can be used, for instance changes of thickness†; or we may resort to any method of graphical integration. Although several different methods exist,‡ all are based upon the equations of equilibrium and involve considerable time. A new method, which has now been worked out by the writer, based entirely upon *elastic equations*, enables us to trace those lines, along which the sum of the principal stresses has a constant value. Since we can assign to each line its appropriate parameter, we know in this way the sum of the principal stresses at any point of the plate. This method is very accurate and is carried out quite simply as follows.

---

* In charge of the photo-elastic investigation in the Mech. Techn. Labor., Technical University, Munich under the direction of Professor L. Föppl.

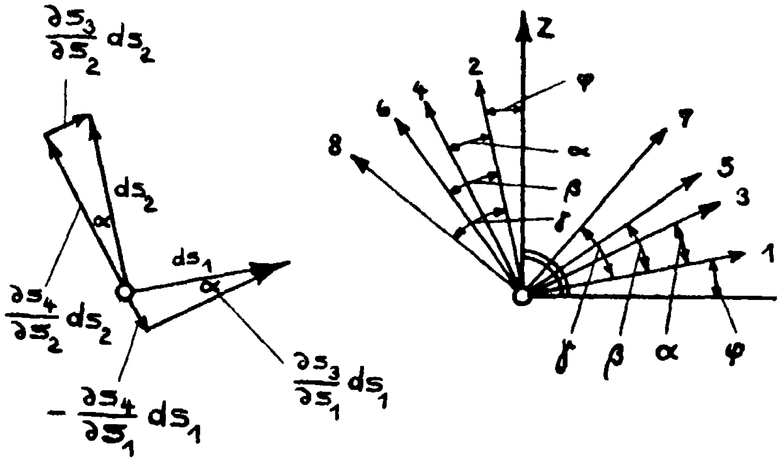
† Coker and Filon, "A Treatise on Photo-Elasticity," Camb., 1931, §§ 2·47-2·49.

‡ Coker and Filon, *loc. cit.*, § 2.30; Filon, 'Brit. Ass.,' Rep. 1923, p. 350; L. Föppl, 'SitzBer. bayer. Akad. Wiss.,' p. 246 (1928); Widdern and Kurzhals, 'Mitt. mech. techn. Lab., München,' vol. 34, p. 4 (1930) and vol. 35, p. 14 (1931); Baud, 'J. Frankl. Inst.,' vol. 211, p. 457 (1931).

(I) THEORY.

(1) *General Relations.*

The lines of principal stress 1 and 2, fig. 1, give at all points the directions of the two principal stresses (σ_1 and σ_2); where ϕ is the angle between the directions 1 and Y, or 2 and Z respectively (Y — Z are Cartesian co-ordinates).



$\frac{\partial s_3}{\partial s_2} ds_2$
ds_2
α
ds_1
α
$\frac{\partial s_4}{\partial s_2} ds_2$
$-\frac{\partial s_4}{\partial s_1} ds_1$
$\frac{\partial s_3}{\partial s_1} ds_1$
Z
2
4
6
8
φ
α
β
γ
7
5
3
1
γ
β
α
φ

FIG. 1.

According to Coker and Filon (*loc. cit.*) the lines, along which the sum of the principal stresses ($\sigma_1 + \sigma_2 = p$) has a constant value, are called *isopachic lines*; and are denoted by the integer 3 in fig. 1, where 4 denotes the lines, which are orthogonal to the isopachic lines, α is the angle between the positive directions 3 and 1 (or 4 and 2 respectively). Since p is constant along the isopachic lines, we have $\partial p/\partial s_3 = 0$, and it therefore follows that

$$\frac{\partial p}{\partial s_1} = \frac{\partial p}{\partial s_4}\frac{\partial s_4}{\partial s_1} = -\sin\alpha\frac{\partial p}{\partial s_4}, \qquad \frac{\partial p}{\partial s_2} = \frac{\partial p}{\partial s_4}\frac{\partial s_4}{\partial s_2} = \cos\alpha\frac{\partial p}{\partial s_4}. \qquad (1)$$

Corresponding equations exist for the *isochromatic lines*, along which the difference of the principal stresses ($\sigma_1 - \sigma_2 = q$) has an equal value, and denoted here by the integer 5, 6 being the corresponding integer for the orthogonal lines; β is the angle between the positive directions 5 and 1 (6 and 2). We obtain

$$\frac{\partial q}{\partial s_1} = -\sin\beta\,\frac{\partial q}{\partial s_6}, \qquad \frac{\partial q}{\partial s_2} = \cos\beta\,\frac{\partial q}{\partial s_6}. \tag{2}$$

Finally, the corresponding equations can be derived from the fact, that along the *isoclinic lines* the angle ϕ is constant. These are denoted by the integer 7, and the orthogonal lines by 8, and γ is the angle between the positive directions 7 and 1 (8 and 2). We find

$$\frac{\partial \phi}{\partial s_1} = -\sin\gamma\,\frac{\partial \phi}{\partial s_8}, \qquad \frac{\partial \phi}{\partial s_2} = \cos\gamma\,\frac{\partial \phi}{\partial s_8}. \tag{3}$$

To obtain the positive directions 4, 6 and 8 let the gradients $\partial p/\partial s_4$, $\partial q/\partial s_6$ and $\partial\phi/\partial s_8$ be supposed positive. From the directions 2, 4, 6, 8 the directions 1, 3, 5, 7 are obtained by clockwise rotation.

(2) *Equations of Equilibrium.*

Formulating the equilibrium conditions for the curvilinear element of fig. 2 we obtain

$$\frac{\partial \sigma_1}{\partial s_1} + (\sigma_1 - \sigma_2)\frac{\partial \phi}{\partial s_2} = 0, \qquad \frac{\partial \sigma_2}{\partial s_2} + (\sigma_1 - \sigma_2)\frac{\partial \phi}{\partial s_1} = 0.$$

These equations can in accordance with

$$\sigma_1 = \frac{p+q}{2}, \qquad \sigma_2 = \frac{p-q}{2}, \tag{4}$$

be put into the form

$$\frac{\partial p}{\partial s_1} + \frac{\partial q}{\partial s_1} + 2q\frac{\partial \phi}{\partial s_2} = 0, \qquad \frac{\partial p}{\partial s_2} - \frac{\partial q}{\partial s_2} + 2q\frac{\partial \phi}{\partial s_1} = 0. \tag{5}$$

Referring back to equations (1), (2) and (3) we now obtain

$$\frac{\partial p}{\partial s_4}\sin\alpha = -\frac{\partial q}{\partial s_6}\sin\beta + 2q\frac{\partial \phi}{\partial s_8}\cos\gamma, \quad \frac{\partial p}{\partial s_4}\cos\alpha = \frac{\partial q}{\partial s_6}\cos\beta + 2q\frac{\partial \phi}{\partial s_8}\sin\gamma. \tag{6}$$

By means of these equations the directions and distances of the isopachics can be derived from those of the isochromatics and isoclinics.

Over the free boundary, which is a line of principal stress, either σ_1 or σ_2 vanishes. If it is a line 2, σ_1 vanishes, and we have

$$p = -q, \qquad \frac{\partial p}{\partial s_2} = -\frac{\partial q}{\partial s_2}.$$

Using the second equation (5) we have*

$$-\frac{\partial p}{\partial s_2} = \frac{\partial q}{\partial s_2} = q\frac{\partial \phi}{\partial s_1}. \tag{7}$$

Dividing the first equation (5) by equation (7), we obtain

$$-\frac{\dfrac{\partial p}{\partial s_1}}{\dfrac{\partial p}{\partial s_2}} + \frac{\dfrac{\partial q}{\partial s_1}}{\dfrac{\partial q}{\partial s_2}} + 2\frac{\dfrac{\partial \phi}{\partial s_2}}{\dfrac{\partial \phi}{\partial s_1}} = 0.$$

Returning to equations (1), (2) and (3) we can write

$$tg\,\alpha = tg\,\beta + 2\cot\gamma. \tag{8}$$

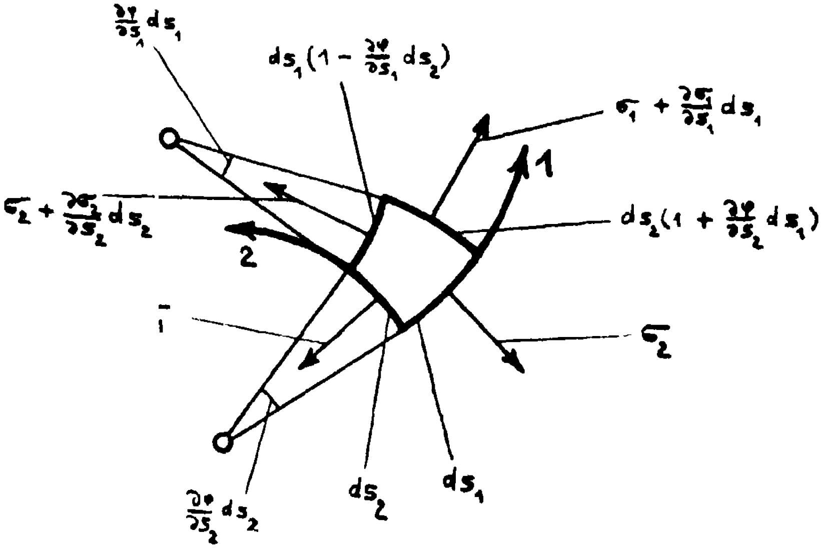

$\frac{\partial \varphi}{\partial s_1} ds_1$
$ds_1(1 - \frac{\partial \varphi}{\partial s_1} ds_2)$
$\sigma_1 + \frac{\partial \sigma_1}{\partial s_1} ds_1$
1
$\sigma_2 + \frac{\partial \sigma_2}{\partial s_2} ds_2$
$ds_2(1 + \frac{\partial \varphi}{\partial s_2} ds_1)$
2
$\bar{1}$
σ_2
$\frac{\partial \varphi}{\partial s_2} ds_2$
ds_2
ds_1

Fig. 2.

If the free boundary is a line 1, σ_2 vanishes, and we obtain :

$$\cot\alpha = \cot\beta + 2\,tg\,\gamma. \tag{9}$$

Therefore at the free boundary the directions of the isopachics are independent of the gradients of q and ϕ.

---

* Equation (7) gives immediately $q = q_0 e^{\int_0^{s_2} \frac{\partial \phi}{\partial s_1} ds_2} = q_0 e^{-\int_0^{s_2} tg\gamma\, d\phi}$. This integral over the free boundary was given originally by L. Föppl.

(3) *Stress—Strain Equations.*

In Cartesian co-ordinates Y — Z with normal stresses σ_y and σ_z, a shear τ_{yz} and displacements η and μ, the stress-strain relations can be put into the form

$$\mathrm{E}\frac{\partial\eta}{\partial y}=\sigma_y-(1/m)\,\sigma_z,\quad \mathrm{E}\frac{\partial\mu}{\partial z}=\sigma_z-(1/m)\,\sigma_y,\quad \tfrac{1}{2}\left(\frac{\partial\mu}{\partial y}+\frac{\partial\eta}{\partial z}\right)=\frac{1+1/m}{\mathrm{E}}\,\tau_{yz},$$

where E is Young's Modulus and $1/m$ is Poisson's ratio.

If we introduce $\sigma_y+\sigma_z=\sigma_1+\sigma_2=p$ and the "rigid body rotation"

$$\frac{\mathrm{K}}{\mathrm{E}}=\tfrac{1}{2}\left(\frac{\partial\mu}{\partial y}-\frac{\partial\eta}{\partial z}\right) \tag{10}$$

we can write

$$(1+1/m)\,\sigma_z+\mathrm{E}\frac{\partial\eta}{\partial y}=p,\quad (1+1/m)\,\tau_{yz}-\mathrm{E}\frac{\partial\eta}{\partial z}=\mathrm{K}.$$

Differentiating with regard to z and y and adding, we have

$$(1+1/m)\left(\frac{\partial\sigma_z}{\partial z}+\frac{\partial\tau_{yz}}{\partial y}\right)=\frac{\partial p}{\partial z}+\frac{\partial\mathrm{K}}{\partial y}.$$

The left side represents the condition of equilibrium in the direction of the Z-axis and therefore vanishes. We find

$$\frac{\partial p}{\partial z}=-\frac{\partial\mathrm{K}}{\partial y}. \tag{11}$$

Eliminating the displacement μ we obtain

$$\frac{\partial p}{\partial y}=\frac{\partial\mathrm{K}}{\partial z}. \tag{12}$$

Therefore p and K are conjugate functions.*

Since equations (11) and (12) are independent of choice of axes, we have along the lines 3 and 4

$$\frac{\partial p}{\partial s_3}=\frac{\partial\mathrm{K}}{\partial s_4},\quad \frac{\partial p}{\partial s_4}=-\frac{\partial\mathrm{K}}{\partial s_3}. \tag{13}$$

Referring to $\partial p/\partial s_3=0$, it follows, that K must be constant along the lines 4. Further the gradients of p and —K are the same. If we trace the lines 3 and 4 with the same intervals ($\Delta p=-\Delta\mathrm{K}$), it follows, that approximately they must form a network with small squares.

---

* This result is already known; see for instance, A. and L. Föppl, "Drang und Zwang," vol. 1, 2nd ed., Munchen und Berlin, 1924, § 43.

(II) THE GRAPHICAL METHOD.

(1) *The Starting Points of the Isopachics at the Boundary.*

The preparatory step in constructing the isopachics consists of determining their starting points at the boundary.

By observations with polarized light the isochromatics are obtained with the successive parameters (interval Δq): $q = 0, \Delta q, 2\Delta q, \ldots$. We now trace those isopachics, whose parameters proceed by the same equal steps, so that

$$\Delta p = \Delta q \tag{14}$$

and $p = 0, \Delta q, 2\Delta q, \ldots, -\Delta q, -2\Delta q, \ldots$.

At the *free boundary* we have $p = q$, if it is a line 1, or $p = -q$, if it is a line 2. Therefore isopachics and isochromatics meet the free boundary at the same points. In this way we obtain a large number of starting points.

At the *loaded boundary* the determination of the starting points is not complicated. Formulating the equilibrium condition of an element at the boundary (f_n may be the normal component of the loading forces, which can be measured, and ε the angle between the direction 1 and the tangent to the boundary) we have

$$f_n = \sigma_1 \sin^2 \varepsilon + \sigma_2 \cos^2 \varepsilon$$

or, using (4)

$$p = 2f_n + q \cos 2\varepsilon. \tag{15}$$

The starting points are preferably those, at which p has one of the values:

$$0,\ \Delta q,\ 2\Delta q,\ \ldots,\ -\Delta q,\ \ldots.$$

(2) *The Directions of the Isopachics.*

To trace the isopachics it is necessary to know their directions at a sufficient number of points.

Let a be the distance between two near isopachics, b that between two near isochromatics and c that between two near isoclinics as that a, b and c are the distances in the directions 4, 6 and 8 respectively; the interval $\Delta\phi$ between two near isoclinics can be supposed constant, generally 5°. Then we can write equations (6) approximately in the form

$$\frac{\Delta p}{a} \sin \alpha = -\frac{\Delta q}{b} \sin \beta + 2q \frac{\Delta \phi}{c} \cos \gamma, \quad \frac{\Delta p}{a} \cos \alpha = \frac{\Delta q}{b} \cos \beta + 2q \frac{\Delta \phi}{c} \sin \gamma.$$

Now let

$$2\frac{q}{\Delta q}\Delta\phi\, b = b', \tag{16}$$

$$\frac{bc}{a} = a', \tag{17}$$

and using (14) we find

$$\sin\alpha = \frac{b'\cos\gamma - c\sin\beta}{a'}, \tag{18}$$

$$\cos\alpha = \frac{b'\sin\gamma + c\cos\beta}{a'}. \tag{19}$$

By means of these equations a graphical construction can be carried out in the following simple manner.

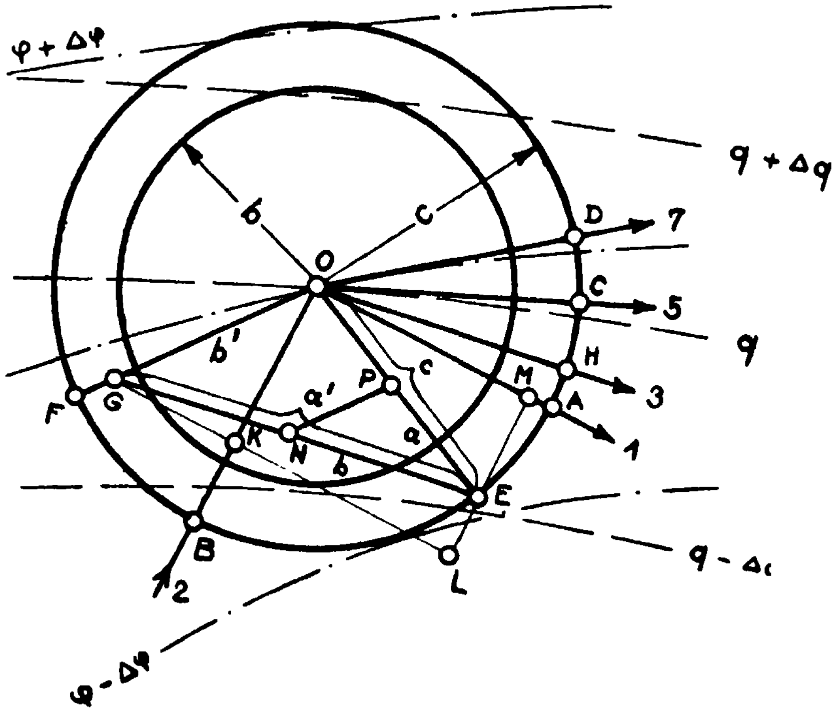


FIG. 3.

At any point O, fig. 3, where an isochromatic intersects an isoclinic, we obtain c by measuring the middle distance between the isoclinic through O of parameter ϕ, and the two near isoclinics, $\phi + \Delta\phi$ at one side and $\phi - \Delta\phi$ at the other side. Correspondingly we obtain b. The length of b', see equation (16), is obtained by means of a nomographical diagram. Fig. 4 represents such a diagram with $\Delta\phi = 5°$ or arc $\Delta\phi = 0{\cdot}08727$. For any abscissa b

the diagram gives us b' as an ordinate with reference to any parameter q. Next we draw the positive direction 1 where the angle ϕ is given by the parameter ϕ of the isoclinic through O, the negative direction 2 and the positive tangents to the isochromatic and the isoclinic. A circle with its centre at O and the radius c intersects these directions at A, B, C and D. Make the arcs AC = AE and AD = BF. Along OF make OG = b' and draw GE. *Then a parallel to* GE *through* O *gives the direction* 3 *at* O.

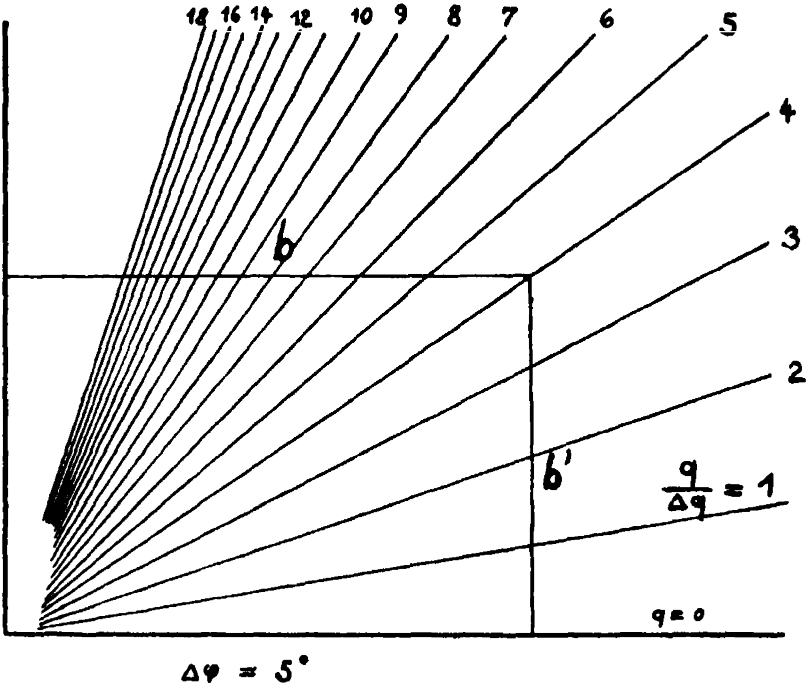


FIG. 4.

To make this construction more intelligible, draw two lines parallel to the directions 1 and 2 through E and G, giving the intersections: K, L and M. Referring to (18) and (19), we have

$$EG = a' \quad , \quad \text{angle } LGE = \alpha.$$

Since GL has the direction 1, it follows that GE gives us the direction 3.

Particular Cases.

(*a*) $b' \gg c : \alpha = 90° - \gamma, \qquad b' = a'$

(*b*) $b' \ll c : \alpha = -\beta, \qquad c = a'$

(*c*) $b' = c,\ \beta = 90° + \gamma$. In this case the point O is an isotropic point of the p — K-network.

Having carried out this construction in a sufficient number of points, the isopachics can easily be drawn, starting at the boundary.

A first check gives each isopachic, which meets the boundary twice. At both starting points the value of p must be the same.

A second check gives the free boundary, where the directions of the isopachics can be derived by means of another construction.

(3) *The Directions of the Isopachics at the Free Boundary.*

By means of equations (8) and (9) the directions of the isopachics at the free boundary can be derived in a very simple manner as shown in fig. 5.

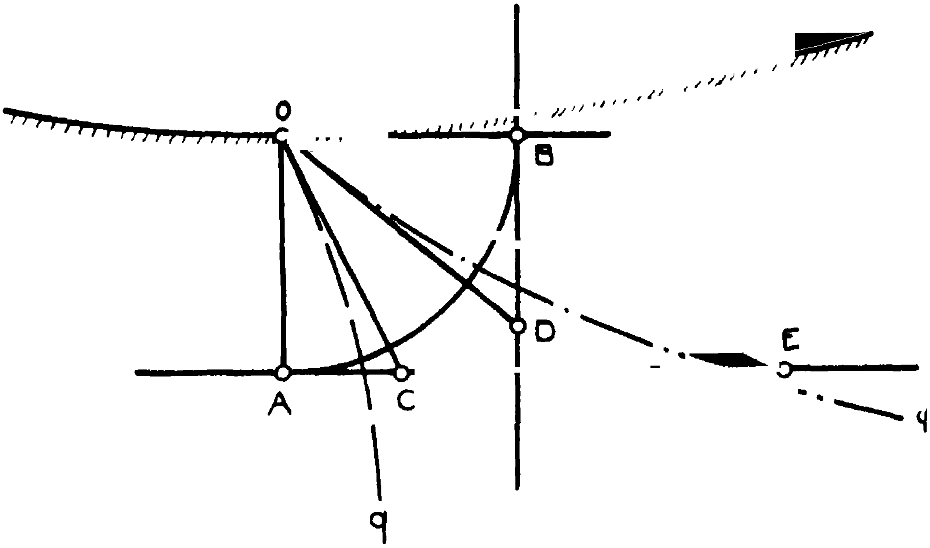


FIG. 5.

At any boundary point O draw the normal and the tangent to the boundary. A circle with O as centre intersects these lines at A and B respectively. The tangent to the isochromatic through O intersects the parallel to the boundary tangent through A at C. The tangent to the isoclinic through O intersects the parallel to the boundary normal through B at D. If no isochromatic or no isoclinic goes through O, the directions 5 or 7 can be determined by interpolation between the isochromatics or isoclinics, which meet the boundary near O. Make CE $= 2$BD, then OE *gives the direction of the isopachic.*

By this construction either equation (8) or (9) is satisfied. If D coincides with B, CE is negative and E lies at the left side of C.

Particular Cases.

(*a*) The boundary is a straight line and therefore an isoclinic; the isopachic and isochromatic then have the same directions.

(*b*) The isoclinic intersects the boundary at right angle; the isopachic then lies in the boundary.

(*c*) The isochromatic lies in the boundary; in this latter case the isopachic lies in the boundary.

The third check ensures that the distance a can be obtained exactly by means of the following construction.

(4) *Exact Determination of the Distance between Two Near Isopachics.*

Referring back to construction II (2), we can derive the intercept a as follows, see fig. 3.

Along EG make EN $= b$ and draw through N a parallel to OG, intersecting OE at P. *Then* EP *gives us the intercept a.*

It is clear, from this construction that equation (17) is satisfied.

Particular Cases.

(*a*) $b' \geqslant c : a = \dfrac{c}{2\dfrac{q}{\Delta q}\Delta\phi}$

In this case the intercept a is obtained immediately by means of the nomographical diagram, fig. 4. With c as ordinate the diagram gives a as abscissa.

(*b*) $b' \leqslant c : a = b.$

In this case the distance between two near isopachics is equal to the distance between two near isochromatics.

It is not necessary to make use of this construction at all points where the directions are derived, but only at those where a check is desired.

(5) *The p — K-network.*

The fourth check is the most important, from the fact, that the isopachics must form with their orthogonal lines a network of small squares. This offers a valuable means of confirming the experimental and graphical work. Further, we obtain in this way the parameters of those isopachics, which do not meet the boundary.

Using all these constructions the isopachics can be obtained very accurately in a short time.

The isopachics, isochromatics and isoclinics together give the complete stress-field analysis, and we obtain immediately values and directions of the two principal stresses at any point.

This graphical method has already been used in many cases and figs. 6 and 7 represent its application to a symmetrical angle plate under pure bending moment.

Summary.

The writer gives a new method of deriving the two principal stresses from the isochromatics and isoclinics. This method enables us to trace the so-called "isopachic lines," along which the sum of the two principal stresses has an equal value. Since we can assign to each line its appropriate parameter we know in this way the sum of the principal stresses at any point of the plate.

The method is based entirely upon *elastic equations*; therefore it is very accurate, offers several alternative methods of checking, and can be carried out in a very simple manner.

Isopachics, isochromatics and isoclinics together give values and directions of the two principal stresses at any point; therefore they represent the *complete field of stress.*

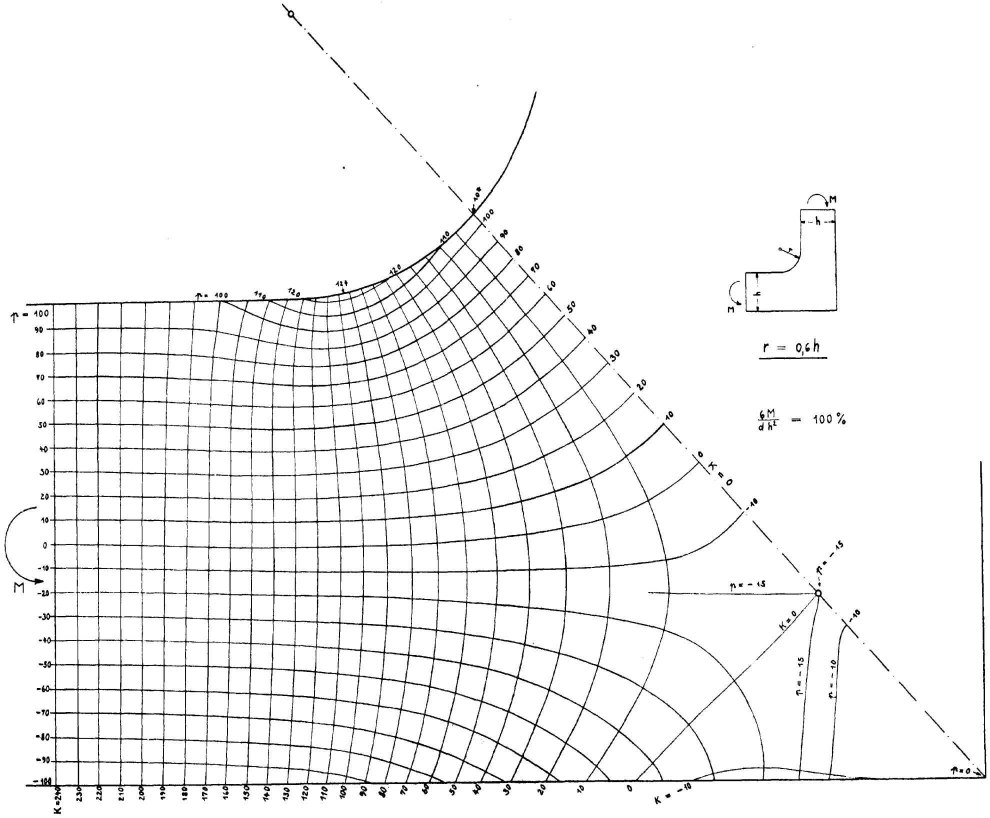


FIG. 6.—p – K-network of a symmetrical angle plate under pure bending moment.

Stress.	Value.	Direction.
σ_1	$\frac{p+q}{2}$	ϕ
σ_2	$\frac{p-q}{2}$	$\phi + 90°$

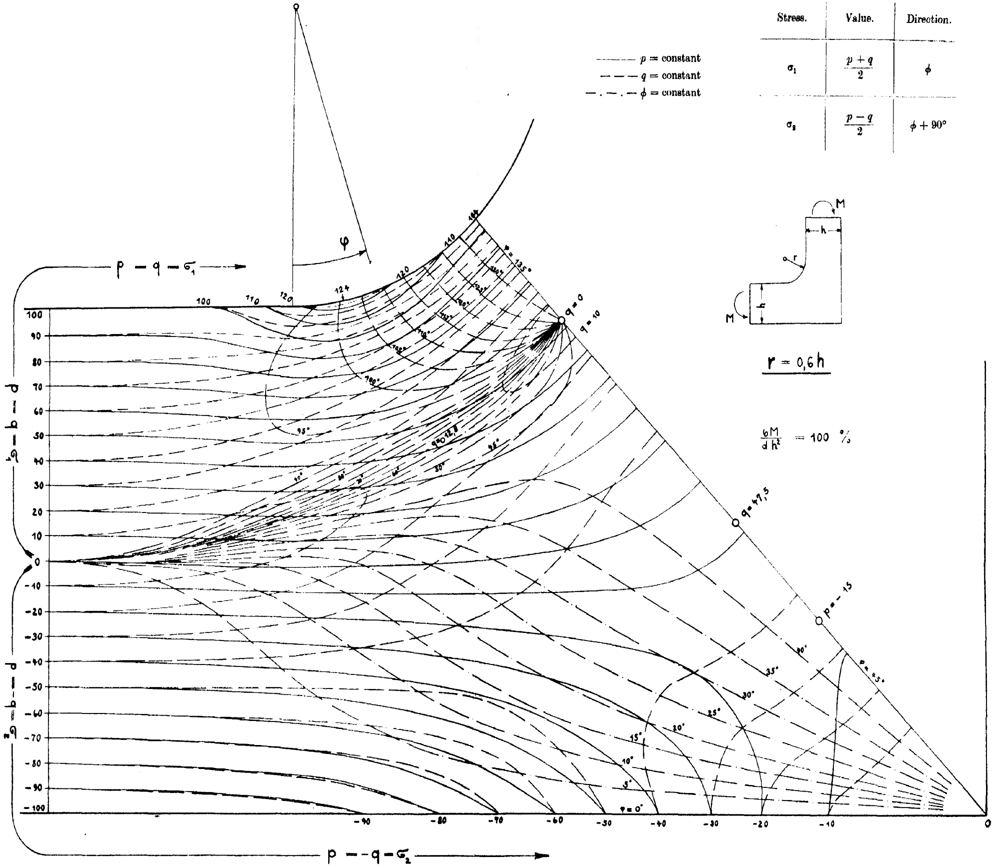


Fig. 7.—Complete stress-field of a symmetrical angle plate under pure bending moment.

On Approximation by Polygons in the Calculus of Variations.

By L. C. Young, M.A., Trinity College, Cambridge.

(Communicated by G. H. Hardy, F.R.S.—Received February 9, 1933.)

1. *Introduction.*

We consider in this note questions of approximation having a bearing on variational problems, although the methods we use have no connection with those of the Calculus of Variations. The hypotheses we make are a good deal simpler than those generally made in variational problems.

The results we obtain include as special cases certain parts of the theory of Weierstrass's E-function, now called the theory of semi-continuity. But the information we obtain is much more precise.

The ideas underlying our methods are elementary. We base our theory on a *generalization of the notion of curve* that suggests itself when we consider the analogous notion of a set of ∞^1 line elements that occurs in the work of *Sophus Lie.* And in point of fact, our generalized curves, although new to pure mathematics, are essentially the same curves as those that occur as paths of particles in modern *quantum mechanics.*

Corresponding to our generalized curves we have also a *generalization of curvilinear integration.* But the new integrals are easy to manipulate and they have a simple connection with the old ones from which they are derived when the integrand is subjected to a process of linear averaging or "folding over" (Faltung) such as occurs in recent work of Norbert Wiener.

2. *Generalized Curves in a Finite Vector Field.*

We consider an interval (t_0, t_1) of the real axis of t and a bounded portion of a Cartesian space of a finite number of dimensions in which the variable will be denoted by x.

By a finite vector field we mean a finite number of vectors θ_i $(i = 1, \ldots, m)$ that are points of x-space.

By an ordinary curve belonging to this field† we mean a polygon $x(t)$ whose rate of increase[†] dx/dt, wherever it exists, is restricted to be one of the vectors θ_i.

---

† When not restricted to belong to a particular field of θ_i an "ordinary curve" will mean a "rectifiable curve" in the ordinary sense. To prevent confusion with generalized curves on the one hand and with polygons on the other we shall write "ordinary rectifiable

Such a polygon is characterized by its first point x_0 together with a set of m monotone increasing functions,

$l_i(t)$ = measure of the parts of (t_0, t) for which $dx/dt = \theta_i$,

and we have the relations

$$\begin{aligned} \sum_{i=1}^{m} l_i(t) &= t - t_0 \\ \sum_{i=1}^{m} \theta_i l_i(t) &= x(t) - x_0 \end{aligned} \tag{i}$$

More generally *we shall agree to say of any set of monotone increasing functions* $l_i(t)$ *given in* (t_0, t_1) *and satisfying the first condition in* (i) *that they define a generalized curve of the field of the* θ_i *commencing at the given point* x_0.

The term generalized curve is here merely a name that we choose to give. We shall call *moveable point* on such a "curve" the quantity $x(t)$ defined by the second of the equations (i).

There is then a formal analogy between an ordinary curve in the field of θ_i and a generalized curve in this field, since in both cases the equations (i) are fulfilled.

On the other hand it would be false to say that the rate of increase of the moveable point $x(t)$ is in any way to be equated to one of the θ_i for a generalized curve. If we wished to do anything of this kind we should have to agree to interpret the words "rate of increase" in a new way altogether. A natural interpretation, that would be in accordance with quantum theory[‡] ideas, would be to regard dx/dt as a symbol capable of m different values θ_i to each of which a definite probability dl_i/dt would be attached.

In virtue of the first equation (i) each l_i is absolutely continuous so that for almost all t the probabilities dl_i/dt would all exist and their sum would be unity.

2.1. The definition of generalized curve in a finite vector field can also be given for *curves defined parametrically*. The difference is that there is no fixed t-axis, and the first of the equations (i) is not longer necessary. We pass to the parametric curve by identifying our original t-axis with a component of the vector x rather than with the parameter that we now denote by the symbol t.

[†] We are considering here a curve defined with reference to a fixed t-axis, that is to say, a definite vector function of t, which has therefore a definite rate of increase. In the next paragraph we shall pass on to the more interesting case of curves defined parametrically with which we shall be concerned all through the remainder of the present paper.

[‡] *Cf.* Dirac, "Quantum Mechanics," p. 31, lines 29–31, and similar references.

It is convenient to make the following agreement with regard to change of variable. Let $f(t)$ and $\phi(u)$ be two monotone increasing functions. By $f(\phi(u))$ we shall then mean *not* the function obtained by substituting $\phi(u)$ for t in $f(t)$, but a slightly different function equal to the latter whenever the value $t = \phi(u)$ is a point of *continuity* of $f(t)$, and also whenever the value $t = \phi(u)$ is assumed by ϕ solely at one point u. And we complete our definition by stipulating that if t_0 is assumed by $\phi(u)$ in the interval (u_0, u_1) then $f[\phi(u)]$ is to be defined in this interval (which may be either open or closed) by linear interpolation between the values $f(t_0 - 0)$ and $f(t_0 + 0)$. This convention is a useful one to make when we change parameters along a curve, whether ordinary or generalized.

2.11. We say that *m real functions $l_i(t)$ monotone increasing in an interval (t_0, t_1) define a generalized parametric curve* C *in the field of vectors θ_i with a given initial point x_0, if the moveable point $x(t)$ defined by the equation*

$$\sum_{i=1}^{m} l_i(t)\,\theta_i = x(t) - x_0, \tag{ii}$$

is a continuous function of t.

And we add the stipulation that the same curve C will be defined by a new set of functions $l_i(\phi(u))$ provided that $\phi(u)$ is monotone increasing and that any discontinuities of ϕ correspond to stretches of constancy of the function $x(t)$ just as with the change of parameter along an ordinary curve.

2.12. We attach no condition of continuity to the functions $l_i(t)$ themselves. In this respect there is a difference between the parametric and the non-parametric case, for we saw that in the non-parametric case the functions $l_i(t)$ had to be absolutely continuous.

We can, however, easily arrange that the same should now hold if we make a change of parameter. We write

$$l(t) = \sum_{i=1}^{m} |\theta_i|\, l_i(t),$$

and $\phi(u)$ for the inverse function of $l(t)$. The functions

$$l_i(\phi(u)),$$

will then be absolutely continuous since their increments are majorized by constant multiples of those of $l[\phi(u)] \equiv u$. On the other hand we have for every pair of values τ, τ' of t

$$|x(\tau) - x(\tau')| \leqslant |l(\tau) - l(\tau')|,$$

on account of (ii). This shows that the parameter u increases at least as fast as the length of arc described by the moveable point $x(t)$. The discontinuities of $\phi(u)$ which correspond to stretches of constancy of $l(t)$ therefore correspond to stretches of constancy of $x(t)$ and u is an admissible parameter.

We can make the analogy with the non-parametric curve complete by supposing that the vectors θ_i are *unit vectors* and that the parameter t is the standard parameter u just defined. Then the relations

$$\left.\begin{aligned} \sum_{i=1}^{m} l_i(t) &= t - t_0 \\ \sum_{i=1}^{m} \theta_i l_i(t) &= x(t) - x_0 \end{aligned}\right\}$$

will again be fulfilled.

3. *Curvilinear Integration Relative to a Finite Vector Field.*

Let $F(x, y)$ denote a continuous function of the pair of vectors x, y satisfying the condition of *homogeneity*

$$F(x, \sigma y) = \sigma F(x, y),$$

for every *positive* scalar σ.

The function F is not restricted to be itself a scalar, it is convenient to allow it to be a vector having an arbitrary finite number of components.

We shall denote by $M(F)$ the maximum of $|F|$ for $|y| = 1$ and by $\varpi(F, \delta)$ the upper bound of the difference $|F(x_1, y) - F(x_2, y)|$ subject to the conditions $|y| = 1$, $|x_1 - x_2| < \delta$, the x's being in both cases restricted to lie in a fixed bounded region.

3.1. We shall consider here only the parametric case of curvilinear integration. Let C be a generalized curve in the field of vectors θ_i, let $l_i(t)$ be the functions by which C is defined and $x(t)$ the moveable point on C. We define

$$I(C, F) = \int_{t_0}^{t_1} \sum_{i=1}^{m} F(x(t), \theta_i)\, dl_i(t),$$

as the curvilinear integral along C of our function F. From the form of this integral as a Stieltjēs integral in the elementary sense it follows at once that it is independent of the parameter t; furthermore, since F satisfies the condition of homogeneity the integral remains unaltered when one or more of the θ_i are multiplied by a positive scalar and at the same time the corresponding functions $l_i(t)$ are multiplied by the inverse of this scalar.

We may therefore suppose without loss of generality that the θ_i are unit vectors and that the functions $l_i(t)$ are absolutely continuous functions satisfying the condition $\sum_{i=1}^{m} l_i(t) = t - t_0$.

An important special case of our curvilinear integral is the quantity

$$L(C) = I(C, |y|) \equiv \Sigma\, |\theta_i| \int_{t_0}^{t_1} dl_i(t),$$

which reduces to $(t_1 - t_0)$ with our special choice of the t-axis. We call this quantity the *length* of our generalized curve C. Similarly $t - t_0$ is the length of arc along our generalized curve, and when we make this special choice of parameter we say that C is referred to its length of arc as parameter.

3.2. *Evaluation of a Difference.*—We consider two generalized curves C and C* in the field of vectors θ_i, $(i = 1, \ldots, m)$ and refer them to the same parameter t in $(0, 1)$, writing $l_i(t)$ and $x(t)$ for the functions characterizing C and $l_i^*(t)$, $x^*(t)$ for those characterizing C*.

We consider a sub-division S of the t-axis into portions of the form (τ_0, τ_1) and we write δl_i, δl^*_i for the increments of $l_i(t)$, $l_i(t)^*$ over such a portion (τ_0, τ_1). We denote by V (S, C − C*), or more shortly V, the quantity

$$|x^*(0) - x(0)| + \sum_{(\tau_0, \tau_1)} \sum_{i=1}^{m} |\delta l^*_i - \delta l_i|,$$

We denote further by L the length $L(C) \equiv \sum_{i=1}^{m} |l_i(1)|$ and by ΔL the length of the longest arc of C determined by the sub-division S, *i.e.*, the greatest of the quantities

$$\sum_{i=1}^{m} |l_i(\tau_1) - l_i(\tau_0)|.$$

3.21. Let t and t^* belong to the same portion (τ_0, τ_1). Then

$$|x^*(t)^* - x(t)| < V + \Delta L.$$

To see this we remark that the expression

$$|x^*(\tau_0) - x(\tau_0)| \leqslant |x^*(0) - x(0)| + \sum^{m} |l^*_i(\tau_0) - l_i(\tau_0)|,$$

is certainly not greater† than the part of the expression for V that refers to the points of division anterior to τ_0, so that it will suffice to show that

$$|x^*(t^*) - x(t)| \leqslant |x^*(\tau_0) - x(\tau_0)| + \sum_{i=1}^{m} |\delta l_i - \delta l^*_i| + \Delta L,$$

† Observing that $l_i(\tau_0) = l_i(\tau_0) - l_i(0)$ is the appropriate sum of δl_i and that $l^*_i(\tau_0)$ is likewise the appropriate sum of δl^*_i.

and in proving this we may suppose, to simplify the notation† that $\tau_0 = 0$, so that $l_i(\tau_0) = l^*_i(\tau_0) = 0$.

We have then

$$|x^*(t^*) - x(t)| \leqslant |x^*(\tau_0) - x(\tau_0)| + \sum_{i=1}^{m} |l^*_i(t^*) - l_i(t)|$$

$$\leqslant |x^*(\tau_0) - x(\tau_0)| + \Sigma_+ l^*_i(\tau_1) + \Sigma_- l_i(\tau_1),$$

where Σ_+ refers to the indices i for which $l^*_i(t^*) > l_i(t)$ and Σ_- to the remaining indices

$$\leqslant |x^*(\tau_0) - x(\tau_0)| + \Sigma_+ |l^*_i(\tau_1) - l_i(\tau_1)| + \sum^{m} l_i(\tau_1)$$

$$\leqslant |x^*(\tau_0) - x(\tau_0)| + \sum_{i=1}^{m} |l^*_i(\tau_1) - l_i(\tau_1)| + \sum_{i=1}^{m} l_i(\tau_1),$$

which, since $l_i(\tau_0) = l^*_i(\tau_0) = 0$, gives us that we wished to prove.

3.22. From what we have just proved we derive the following result

$$\left| \int_{\tau_0}^{\tau_1} F(x(t), \theta_i)\, dl_i(t) - \int_{\tau_0}^{\tau_1} F(x^*(t), \theta_i)\, dl^*_i(t) \right|$$

$$< \varpi(F, V + \Delta L) \,.\, \delta l_i + M(F)\, |\delta l^*_i - \delta l_i|.$$

We obtain this by noticing that if $\delta l_i^* = 0$ one integral vanishes and the other has modulus less than $M(F)\, \delta l_i$ which is the same as $M(F)\, |\delta l^*_i - \delta l_i|$; while if $\delta l^*_i > 0$ we can write the difference whose modulus we wish to evaluate in the form

$$\frac{1}{\delta l^*_i} \int_{\tau_0}^{\tau_1} \int_{\tau_0}^{\tau_1} \{F(x(t), \theta_i) - F(x^*(t^*), \theta_i)\}\, dl_i(t)\, dl^*_i(t^*)$$

$$+ (\delta l_i - \delta l^*_i) \frac{1}{\delta l^*_i} \int_{\tau_0}^{\tau_1} F(x^*(t^*), \theta_i)\, dl^*_i(t).$$

3.23. The inequality of 3.22 gives us by addition the following *evaluation of the difference of curvilinear integrals* along C and C*,

$$|I(C^*, F) - I(C, F)| < L \,.\, \varpi(F, V + \Delta L) + V \,.\, M(F).$$

This result forms the basis of our approximation theorems.

---

† The inequality we are aiming at involves only differences and is therefore unaffected by subtraction of constants from $l_i(t)$ and $l^*_i(t^*)$.

4. *Fundamental Approximation Theorems.*

THEOREM I.—*Let* C^* *be a generalized curve in the field of a finite number of vectors* θ_i *and let the moveable point of* C^* *describe a rectifiable curve* C *in the ordinary sense. Then there is a sequence of polygons* P_n *of bounded lengths converging to* C[†] *such that*[‡]

$$I(P_n, F) \rightarrow I(C^*, F).$$

4.1. The proof of this is immediate. Let the parameter t to which we refer C^* be its length of arc (not to be confused with the length of arc of C) as let $0, L^*$ be its interval of definition.

We divide $0, L^*$ into abutting portions of length less than δ.

We denote by $l^*_i(t)$ and $x^*(t)$ the characteristic functions and the moveable point of C^* and by $l_i(t)$, $x(t)$ those of a polygon P_δ in the field of the θ_i defined in such a manner that

$$l^*_i(\tau) = l_i(\tau),$$

at the points of division τ of the sub-division, and that each portion into which we have divided $0, L^*$ is the sum of a finite number of parts in each of which only one of the functions $l_i(t)$ does not remain constant.

We stipulate further that $x(0) = x^*(0)$.

The two generalized curves C^* and P_δ fulfil the conditions of our lemma 3.23 and we deduce

$$|I(C^*, F) - I(P_\delta, F)| < L^* \cdot \varpi(F, \delta),$$

which tends to zero as $\delta \rightarrow 0$.

P_δ is, however, the same as the ordinary polygon described by its moveable point and the length of P_δ is equal to L^*.

By taking a suitable sequence of numbers δ and the appropriate polygons we have what we require and theorem I is proved.

4.2. THEOREM II.—*Let* P_n *be a sequence of polygons of length less than* K *converging to an ordinary rectifiable curve* C *and suppose that*

$$I(P_n, F),$$

tends to a finite limit a as $n \rightarrow \infty$. *Then there exists a generalized curve* C^*

[†] Convergence of curves is in the sense of Fréchet, 'Rend. Circ. Mat. Palermo,' vol. 22, p. 1 (1906); see also Tonelli, "Fondamenti di Calcolo delle Variazioni," vol. 1, p. 76.

[‡] By the symbol $I(P_n, F)$ where P_n is a polygon we mean the ordinary curvilinear integral $\int F\left(x, \frac{dx}{dt}\right) dt$ taken along P_n.

belonging to a finite vector field[†] *such that the moveable point of* C* *describes the given curve* C *and that*

$$|\mathrm{I}(\mathrm{C}^*, \mathrm{F}) - a| < \varepsilon.$$

4.3. The proof of theorem II is a little longer but quite straightforward, and it depends like that of theorem I on our lemma regarding the evaluation of a difference of integrals.

Let $x_n(t)$ be the moveable point of P_n in terms of a parameter t proportional to the length of arc of P_n and defined in the fixed interval $0 \leqslant t \leqslant 1$.

The differential coefficient $x_n(t)$ which exists except at a countable set of points has modulus $\leqslant \mathrm{K}$. It follows that the functions $x_n(t)$ are equally continuous.[‡]

Since they are bounded for $t = 0$ we can select a subsequence for which $x_n(t)$ tends uniformly to a limit, $x(t)$ say, in $(0, 1)$. The limit $x(t)$ will be, of necessity, a representation of the curve C in $0 \leqslant t \leqslant 1$.

The difference ration $\Delta x/\Delta t$ limit of $\Delta x_n/\Delta t$ will be bounded by the constant K and consequently $x(t)$ will be absolutely continuous and its differential coefficient $x(t)$ will exist for almost all t and will be bounded by K.

4.31. We choose a sub-division S of $(0, 1)$ into intervals of length less than δ/K in such a manner that

$$\frac{\Delta x}{\Delta t} - x \quad dt < \delta/\mathrm{N},$$

where N is the number of co-ordinate axes of x-space and where $\Delta x/\Delta t$ represents a stretchwise constant function of t whose value in any portion τ_0, τ_1 into which S divides $(0, 1)$ is constant and equal to the corresponding difference ratio of $x(t)$.

The integral in question tends to zero as S describes a suitable sequence of sub-divisions (since its integrand does so for almost all t and is less than 2K).

We denote similarly by $\Delta x_n/\Delta t$ a stretchwise constant function of t bearing a similar relation to the difference ratio of $x_n(t)$, and we choose from a uniformly convergent subsequence of P_n, one of them for which

$$|x_n(t) - x(t)| < \delta$$

$$\int_0^1 \left| \frac{\Delta x_n}{\Delta t} \qquad \frac{\Delta x}{\Delta t} \right| dt < \delta/\mathrm{N},$$

and

$$|\mathrm{I}(\mathrm{P}_n, \mathrm{F}) - a| < \tfrac{1}{2}\varepsilon.$$

[†] No particular field of vectors was attached to the polygons P_n.

[‡] Or "equi-continuous" according to Hobson, "Functions of a Real Variable," pp. 167–169.

The second of these conditions can be fulfilled like the other two for any given $\delta > 0$ and $\varepsilon > 0$ since its left-hand side is majorized by the greatest difference $\left.\frac{\Delta x_n}{\Delta t} - \frac{\Delta x}{\Delta t}\right|$ at the finite set of points of division.

The polygon P_n will from now on be fixed, when δ and ε are fixed, and *we shall for simplicity suppose that* $n = 1$

4.32. We denote by θ_h^1 $(h = 1, \ .\ , m)_1$ the unit vectors parallel to the directions of P_1 and we define the functions $l_h^1(t)$ for $h = 1, \ .\ , m_1$ to be the monotone increasing functions that characterize P_1 in this field, that is to say

$l_h^1(t) =$ Length in $(0, t)$ for which P_1 has the direction θ_h^1 $(h = 1, \ldots, m_1)$.

We define a second set of functions $l_h(t)$ (for $h = 1, \ldots, m_1$) also monotone increasing in the following way :—

(i) At the points of division τ of S, $l_h(\tau) = l_h^1(\tau)$.

(ii) In each portion τ_0, τ_1 into which S divides $(0, 1)$ the functions $l_h(t)$ are linear functions.

We shall have

$$x_1(t) - x_1(0) = \sum_{h=1} \theta_h^1 l_h^1(t),$$

while the functions $l_h(t)$ will satisfy the condition

$$\sum_{h=1}^{m_1} \theta_h^1 l_h(t) = \int_0^t \frac{\Delta x_1}{\Delta t}\, dt,$$

in which $\Delta x_1/\Delta t$ refers to the stretchwise constant function already considered.

4.33. We denote by θ_i $(i = 1, \ldots, m)$, $(m \leqslant m_1 + 2N)$ the system of unit vectors θ_h^1 together with the 2N directions parallel or anti-parallel to the coordinate axes. The latter which we may denote by θ_k^2 $(k = 1, \ldots, 2N)$ may for simplicity be supposed not to occur among the θ_h^1, although mere verbal changes would have to be made if this supposition did not hold.

The vectors θ_k^2 $(k = 1, \ldots, 2N)$ form a base from which every vector in x-space can be expressed uniquely as a sum $\sum_k \alpha_k \theta_k^2$ with positive coefficients α_k as small as possible.† We express in this way the vector

$$\frac{d}{dt}\left\{x(t) - x_0 - \sum_{h=1}^{m_1} \theta'_h l_h(t)\right\} \equiv \dot{x} -$$

---

† The θ_k^2 for which $\alpha_k \neq 0$ then form a linearly independent subset consisting of coordinate directions with a sense such that the projection of the given vector along them is positive. These projections will then be the non-vanishing coefficients α_k.

and we denote by $l_i(t)$ the integral form 0 to t of the θ_k^2 component of this quantity, *whenever the index* $i > m_1$ *refers to the vector* θ_k^2 *in the field of* θ_i.

We shall have

$$x(t) - x(0) - \sum_{i=1}^{m} \theta_i l_i(t) = \left\{ x(t) - x_0 - \sum_{h=1}^{m_1} \theta_h{}^1 l_h(t) \right\}$$

$$- \int_0^t \frac{d}{dt} \left\{ x(t) - x_0 - \sum_{h=1}^{m_1} \theta_h{}^1 l_h(t) \right\} dt = 0.$$

We define further

$$l_i{}^1(t) = 0 \qquad \text{for } i > m_1.$$

4.34. We define C* to be the generalized curve in the field of the

$$\theta_i \ (i = 1, \ldots, m),$$

which has $x(0)$ as its beginning and $l^*_i(t) = l_i(t)$ as its characteristic functions. The moveable point of C_2 will then, on account of the relation,

$$\sum_{i=1}^{m} \theta_i l_i(t) = x(t) - x(0),$$

be precisely $x(t)$ which describes the curve C.

On the other hand the polygon P_1 may be considered to belong to the complete field of θ_i $(i = 1, \ldots, m)$ and its characteristic functions are then the functions $l_i{}^1(t)$ $(i = 1, \ldots, m)$ of which the last 2N are now identically zero.

We have for $i > m_1$

$$l_i(1) \leqslant \int_0^1 x - \frac{\Delta x_1}{\Delta t} \ dt < \delta/N.$$

Hence

$$\sum_{\tau_i \tau_1} \sum_{i=1}^{m} |\delta l_i{}^1 - \delta l^*_i| \equiv \sum_{i=m_1}^{m} l_i(1) < 2N\,\delta/N = 2\delta.$$

Thus $V(S, P_1 - C^*) < 3\delta$, and since $\Delta L < \delta$ we find that the evaluation of 3.2 gives

$$|I(C^*, F) - I(P_1, F)| < K \,.\, \varpi(F, 4\delta) + 2\delta \,.\, M(F)$$

$$< \tfrac{1}{2}\varepsilon,$$

by choice of δ. Consequently we have, in this case,

$$|I(C^*, F) - a| < \varepsilon,$$

which proves theorem II.

5. *Folds and Cycles of a Function* F.

We shall now seek to express our integrals along generalized curves in terms of ordinary curvilinear integrals. In order to do this we introduce the notions of folds and cycles of $F(x, y)$—notions that are analogous to certain linear transforms of functions of a real variable that have been employed recently by Norbert Wiener and others.

5.1. By a *fold*† *of* F in the field of θ_i, we shall mean a function of (x, y) having the form

$$\sum_{i=1}^{m} F[x, p_i(x, y)],$$

where the p_i are a finite set of continuous vector functions of (x, y) with constant directions θ and satisfy the conditions

$$p_i(x, \sigma y) = \sigma p_i(x, y) \qquad \text{for every } \sigma > 0$$

$$\sum_{i=1}^{m} p_i(x, y) = y.$$

By a *cycle of* F at a point x_0 we shall mean a sum of the form

$$\sum_{i=1}^{m} F(x_0, q_i),$$

where the q_i are vectors with the directions θ_i connected by the relation

$$\sum_{i=1}^{m} q_i = 0.$$

5.2. We consider a rectifiable curve C in the ordinary sense and we form the curvilinear integral along C of a fold $\tilde{F}(x, y)$ of the function F. To this expression we now add a finite number of cycles $\kappa(F, x_j)$ of F that refer to points x_j (not necessarily distinct) of the cuve C.

It is easy to see that these cycles can be approximated by integrals along C of cycles varying continuously with the points of C, so that the effect of adding them to the integral of $\tilde{F}$ will not differ appreciably from what we should get by making use of a different fold $\tilde{F}^*$ of F. But this remark, which is, of course, only valid when C does not reduce to a single point, need not concern us.

---

† Substantially what Wiener calls "Faltung" ("The Fourier Integral," p. 45, line 8), although his process involves two functions instead of only one, the reason for this difference being that we have restricted our field to be finite. Another difference is that his process is applied to $f(x - y)$ while ours is applied to $F(x, y)$.

5.21. *We shall show that the sum*

$$\mathrm{I}(\mathrm{C}, \tilde{\mathrm{F}}) + \sum_j \kappa(\mathrm{F}, x_j)\dagger$$

is equal to the integral of F *along a generalized curve* C* *in the field of the* θ_i *such that the moveable point of* C* *describes* C.

To see this we define our m monotone increasing functions $l_j(t)$ vanishing for $t = 0$ to have, at (say) the first value t_j of t for which the moveable point $x(t)$ reaches the value x_j, discontinuities of amount $|q_i|$ where q_i is the vector of direction θ_i that refers to the cycle $\kappa(\mathrm{F}, x_j)$. These discontinuities will have to be added together if several x_j coincide. Between the points t_j we stipulate further that $l_i(t)$ is to agree except for an additive constant with the indefinite integral of $|p_i[x(t), \dot{x}(t)]|$, this quantity being certainly summable if t agrees with the length of arc—and indeed for all admissible parameters on account of homogeneity.

We verify at once that

$$\sum_{i=1}^{m} \int_0^t \mathrm{F}[x(t), \theta_i]\, dl_i(t) = \int_0^t \tilde{\mathrm{F}}(x(t), \dot{x}(t))\, dt + \sum_{t_j < t} \kappa[\mathrm{F}, x(t_j)],$$

and this relation remains true if the function $\mathrm{F}(xy)$ is replaced by y itself, in which case it gives

$$\sum \theta_i l_i(t) = x(t) - x(0),$$

so that the generalized curve C* defined by the $l_i(t)$ has $x(t)$ for its moveable point.

5.22. We shall now show *conversely*, that *if we are given a generalized curve* C* *in the field of the* θ_i *and the moveable point of* C* *describes a curve* C, *then there exist an expression of the form*

$$\mathrm{I}(\mathrm{C}, \tilde{\mathrm{F}}) + \sum_j \kappa(\mathrm{F}, x_j),$$

where the x_j *lie on* C, *approximating as closely as we please to* I (C*, F).

We write the monotone functions by means of which C* is defined in the form‡ $\lambda_i(t) + \mu_i(t)$ where $\lambda_i(t)$ is absolutely continuous and where $\mu_i(t)$ has derivates zero for almost all t.

---

† When C is an ordinary rectifiable curve we mean by $\mathrm{I}(\mathrm{C}, \tilde{\mathrm{F}})$ the ordinary curvilinear integral along C

$$\int \tilde{\mathrm{F}}\left(x, \frac{dx}{dt}\right) dt = \int \sum_{i=1}^{m} \mathrm{F}\left[x, p_i\left(x, \frac{dx}{dt}\right)\right] dt.$$

‡ *Cf.* De la Vallée Poussin, "Integrales de Lebesgue," p. 93.

The sum $\sum_1^m \theta_i \mu_i(t)$ whose derivate are zero for almost all t is absolutely continuous since it can be written in the form $x(t) - x_0 - \sum_{i=1}^m \theta_i \lambda_i(t)$. Consequently it is constant, and so zero.

The expression

$$\sum_{i=1}^m \int F[x(t), \theta_i]\, d\mu_i(t),$$

regarded as a finite sum of elementary Stieltjes integrals can be approximated as closely as we please by a system of Cauchy sums relating to a sub-division independent of the index i, that is to say by an expression of the form

$$\sum_{i=1}^m \sum_j F[x(t_j), \theta_i]\{\mu_i(t_j) - \mu_i(t_{j-1})\},$$

which on inverting the order of summation (and noticing that

$$\sum_{i=1}^m \theta_i\{\mu_i(t_j) - \mu_i(t_{j-1})\},$$

vanishes for each j) reduces to a sum of cycles of F

$$\sum_j \kappa(F, x_j).$$

There now remains only to deal with the absolutely continuous part of our integral, that is with the expression

$$\sum_{i=1}^m \int F[x(t), \theta_i]\, d\lambda_i(t),$$

and we shall show that we can approximate to it by an $I(C, \tilde{F})$. We can verify this in a number of ways, the simplest of which seems to be the following which depends on integration in the sense of Radon.†

We suppose our parameter t to coincide with the length of arc along C so that $\dot{x}(t)$ exists with unit modulus for almost all t. We denote by T(E) the function of the sets E of values of x, y for which $|y| = 1$, defined so as to equal the increment of t over the totality of points of the t-axis for which $\dot{x}(t)$ exists with unit modulus and the pair of vectors $x = x(t)$, $y = \dot{x}(t)$ lies in E.

We denote similarly by $\Lambda_i(E)$ the corresponding increment of $\lambda_i(t)$, and we remark that in virtue of a well-known theorem of Radon and de la Vallée Poussin

$$\Lambda_i(E) = \int_E \phi_i(x, y)\, dT.$$

---

† 'SitzBer. Akad. Wiss. Wien,' vol. 122, p. 1295 (1913).

We shall have also

$$\sum_{i=1}^{m} \theta_i \Lambda_i(E) = \int_E y\, dT,$$

since $\sum_1^m \theta_i \lambda_i(t) = \int_0^t x(t)\, dt$, and it follows that outside T-measure zero

$$\sum^m \theta_i \phi_i(x, y) = y.$$

We shall have further

$$\int F(x(t), \theta_i)\, d\lambda_i(t) = \int F(x, \theta_i)\, d\Lambda_i = \int F(x, \theta_i)\, \phi_i(x, y)\, dT$$

$$= \int F[x(t), \theta_i]\, \phi_i[x(t), \dot{x}(t)]\, dt,$$

and we can approximate to the $\phi_i(x, y)$ by continuous functions for which this integral is altered arbitrarily little and for which at the same time the sum $\sum_1^m \theta_i \phi_i$ becomes equal to y for all (x, y) when $|y| = 1$. Writing now $p_i(x, y) = \theta_i\, \phi_i(x, y)$ when $|y| = 1$ and prolonging the definition of p_i by the condition $p_i(x, \sigma y) = \sigma p_i(x, y)$ for all $\sigma \geqslant 0$, we find that the expression

$$\sum_{i=1}^{m} \int F(x(t), \theta_i)\, d\lambda_i(t),$$

differs arbitrarily little from an integral of the form $I(C, \tilde{F})$. It follows that our generalized integral

$$I(C^*, F) = \int F(x(t), \theta_i)\, d\{\lambda_i(t) + \mu_i(t)\},$$

differs arbitrarily little from an expression of the form

$$I(C, \tilde{F}) + \sum_j \kappa(F, x_j),$$

and this is what we wished to establish.

5.3. We shall now state a few results relative to folds and cycles of F easily deducible from the theorems we have established until now.

5.31. *Let* P_n *be a sequence of polygons of bounded lengths tending to a rectifiable curve* C *and let* $I(P_n, F)$ *tend to a finite limit* a. *Suppose further that on* C *there is a cycle of* F, *or a sum of such cycles, having the value* b. *Then there will exist a sequence of polygons* Q_n *tending to* C *for which* $I(Q_n, F)$ *tends to the limit* $a + b$.

5.32. *Let* $\tilde{F}$ *be a fold of* F *and* P_n *a sequence of polygons of bounded lengths tending to* C, *such that* $I(P_n, \tilde{F})$ *tends to a finite limit* a, *then there will exist a sequence of polygons* Q_n *tending to* C *for which* $I(Q_n, F)$ *tends to* a.

6. *Convexity and Semi-continuity.*

We shall call a real function $F(x, y)$ of the pair of vectors x, y *convex* for a particular pair of values x_0, y_0, if whenever $\eta_1, \ldots \eta_m$ is a system of vectors of sum y_0 we have

$$\sum_{i=1}^{m} F(x_0, \eta_i) \geqslant F(x_0, y_0).$$

The term convex refers here, of course, only to the dependence of F on its second argument.

We shall call $F(x, y)$ *convex along a rectifiable curve* C if—

(i) the values of F are real numbers;
(ii) the cycles of F at all points of C are non-negative;
(iii) at almost all points of C, F is convex for the pair of values $x(t)$, $x(t)$ of x, y.

6.1. It is easy to see that the necessary and sufficient condition for a real F to be convex along C is that for every fold $\widetilde{F}$ of F and for every system of a finite number of cycles of F on C, $\kappa(F, x_j)$ we shall have

$$I(C, F) \leqslant I(C, \widetilde{F}) + \sum_j \kappa(F, x_j).$$

In consequence we may also say that the condition

$$I(C, F) \leqslant I(C^*, F),$$

for every generalized curve whose moveable point described C will also be necessary and sufficient for the convexity of F along C. Hence, in order that for every sequence of polygons P_n of bounded lengths for which P_n tends to C and the quantity $I(P_n, F)$ has a finite limit we shall have

$$\text{Lim.}\ I(P_n, F) \geqslant I(C, F),$$

it is necessary and sufficient that F be convex along C.

6.2. Let us now consider a sequence of rectifiable curves C_n converging to C, and suppose that for a particular way of folding—that is a particular choice of the vectors $p_1(x, y), \ldots, p_m(x, y)$—the quantity $I(C_n, \widetilde{|y|})$ remains bounded.

I say that in this case, if F is convex along C, we shall have

$$\underset{n \to \infty}{\text{Lim. inf.}}\ I(C_n, \widetilde{F}) \geqslant I(C, F).$$

For if this were false we could select a *subsequence* for which, denoting by Φ a vector with the two components F and $|y|$, the quantity

$$I(C_n, \widetilde{\Phi}),$$

would tend to a limiting vector whose first component was less than I (C, F). And by theorems proved in 4 and 5 we could approximate to C_n by a polygon P_n for which I (P_n, Φ) would differ as little as we pleased from I (C_n, $\tilde{\Phi}$). We could therefore select P_n so that $P_n \rightarrow C$ and I (P_n, Φ) had the same limit as I (C_n, $\tilde{\Phi}$). In particular, the lengths of the P_n as components of the quantities I (P_n, Φ) would be bounded, while the limit of I (P_n, F) would exist and be less than I (C, F), which we know not to be the case if F is convex along C.

6.3. We have thus proved that *in order that for every sequence of curves* C_n *of bounded lengths converging to* C *we should have*

$$\operatorname*{Lim.\,inf.}_{n \rightarrow \infty} \; I(C_n, F) \geqslant I(C, F),$$

it is necessary and also sufficient that F *be convex along* C. And this condition is at the same time sufficient to ensure the more general result that if $C_n \rightarrow C$ and I (C_n, $\widetilde{|y|}$) remains bounded $\operatorname*{Lim.\,inf.}_{n \rightarrow \infty}$ I (C_n, $\tilde{F}$) $\geqslant$ I (C, F).

6.4. We have obtained our semi-continuity theorem without any reference to classical criteria of the Calculus of Variations. We now observe that in virtue of classical results of the theory of convex bodies and their planes of support (Stützebenen) it is necessary and sufficient for convexity of F at a point $x_0 y_0$ that there should exist numbers μ_h such that for an arbitrary vector η of components η_h along the axes the expression

$$F(x_0, \eta) - F(x_0, y_0) - \sum_h \mu_h \cdot (\eta_h - y_{0h}),$$

which we write for shortness

$$F(x_0 \eta) - F(x_0 y_0) - \Sigma\, \mu \cdot (\eta - y_0),$$

shall be non-negative.

In particular, choosing for η a vector whose *h*th component alone differs from that of y_0 we have by making $\eta \rightarrow y_0$ from both sides

$$\mu_h \leqslant \text{Lower limit} \left\{ \frac{F(x_0, y_0 + \delta) - F(x_0 y_0)}{|\delta|} \right\}$$

$$\mu_h \geqslant \text{Upper limit} \left\{ \frac{F(x_0, y_0 - \delta) - F(x_0 y_0)}{-|\delta|} \right\},$$

where δ denotes a small vector along the *h*th co-ordinate axis. And if F has partial derivates with respect to the components of y our expression reduces to the E-function of Weierstrass

$$E(x_0, y_0, \eta) = F(x_0 \eta) - F(x_0 y_0) - \Sigma \frac{\partial F}{\partial y_h} \cdot (\eta_h - y_{0h}).$$

If the partial derivates do not exist we may still use the same nomenclature, remembering that there may possibly be a number of different E-functions in such a case.

6.5. We can now restate the semi-continuity theorem in the following manner :—

In order that the function of lines I (C, F) *be lower semi-continuous for* $C = C_0$ *in a neighbourhood consisting of curves* C *of bounded lengths it is necessary and sufficient that—*

(i) *The (real) function* F *should have no negative cycles on* C_0.

(ii) *There exist p.p* † *on* C_0 *a non-negative* E-*function.*

We shall return later to the significance of this and our other results in connection with the Calculus of Variations.

Summary.

In this paper we introduce notions closely related to those of modern quantum mechanics. We give to these notions a rigorous form by means of Stieltjēs integration.

We obtain new methods for analysing the behaviour of a curvilinear integral I_C when C is a variable curve of bounded length and we show how to determine the limits to which I_C will tend when C approximates to a particular curve C_0.

As a special application we consider the case in which the integrand of I_C is a scalar function of the position and velocity and we obtain general conditions for semi-continuity.

Throughout the paper we take our curves to be defined parametrically.

---

† p.p. stands for " presque partout " (almost everywhere).

The Infra-Red Absorption Spectrum of Nitrogen Tetroxide and the Structure of the Molecule.

By G. B. B. M. SUTHERLAND, Trinity College, Cambridge.

(Communicated by T. M. Lowry, F.R S —Received February 17, 1933.)

History.

The earliest observations on the infra-red absorption spectrum of nitrogen peroxide were made by Warburg and Leithauser,* who found that at ordinary temperatures the mixture of the tetroxide and dioxide had strong absorption maxima at 3·4 μ, 5·7 μ, and 6·1 μ. By varying the temperature, and so the degree of dissociation of tetroxide molecules into dioxide molecules, they were able to attribute the absorption band at 5·7 μ to the nitrogen tetroxide molecule, and show that the other two bands were due to the molecule of nitrogen dioxide. A later investigation by Eva von Bahr† confirmed these observations and disclosed another band due to the dioxide at 7·3 μ. None of these investigators examined the region beyond 8 μ, and all the work was done using low dispersion. Recently Strong and Woo‡ have examined the spectrum between 23 μ and 150 μ and find two main regions of absorption, one near 26 μ, and the other around 80 μ.

Experimental.

The spectrometer used in the first survey of the spectrum was a prism instrument having a 60° rock salt prism in the Wadsworth mounting. Subsequently several of the bands were examined on grating instruments of high resolving power. The absorption cell used in the preliminary work was of brass, 20 cm. in length, and with rock salt windows. This cell was lagged with asbestos and could be heated electrically until the temperature of the gas which it contained was as high as 160° C. The absorption cell used in the examination of the individual bands was of glass, 10 cm. long, and with mica windows. For work in the region beyond 7 μ the mica windows were replaced by ones of rock salt. The nitrogen peroxide was taken from a very pure sample specially prepared for research work.* It was kept in a sealed pyrex tube as a liquid (under its own pressure) and the gas was transferred to the cells as required, drying tubes containing phosphorus pentoxide being used to avoid contamination from water vapour in the atmosphere.

---

* 'Ann. Physik,' vol. 23, p. 209 (1907); vol. 28, p. 313 (1909).

† 'Verh. deuts. phys. Ges.,' vol. 15, pp. 673, 710 (1913); 'Ann. Physik,' vol. 33, p. 585 (1910).

‡ 'Phys. Rev.,' vol. 42, p. 267 (1932).

In the first examination of the spectrum, the brass cell was filled with the mixture of gases at room temperature and pressure. The absorption was determined from 1 μ to 14 μ. The cell was then slowly heated until the temperature of the gas was between 150° C. and 160° C., a tap being opened from time to time so that atmospheric pressure was maintained by the escape of some of the heated gas. After equilibrium had been reached at this temperature the tap was kept closed, and the absorption spectrum of the heated gas was found between the same limits.† Finally the gas was pumped out and a run was made with the empty cell. The result of these observations is illustrated in fig. 1. Below 2·8 μ the absorption was apparently continuous. It would appear that the number of overtone and combination bands falling in this region is too great for them to be resolved by this spectrometer, and we hope to reinvestigate it later under high dispersion. The curves in fig. 1 are merely meant to indicate roughly the positions and relative intensities of the various bands; there is, however, no relation between the intensity of a band in fig. 1A and the intensity of the same band in fig. 1B or fig. 1C. The curves represent the composite result of several sets of observations with different amounts of gas in the cell.

It will be seen that in addition to the bands observed by Warburg and Leithauser at 3·4 μ, 5·7 μ, and 6·1 μ, several new bands were observed of which the strongest lie at 7·9 μ and 13·3 μ. It would appear that the band observed by Eva von Bahr at 7·3 μ may be spurious and due to the sodium nitrite formed when the nitrogen peroxide attacks the rock salt windows. It is interesting to note that this band at 7·3 μ and the one at 12 μ, which is also due to the deposit on the windows of the cell, agree fairly closely with Raman lines found in crystalline sodium nitrite.‡ The method employed here to separate the spectrum of nitrogen tetroxide from that of nitrogen dioxide is not entirely satisfactory, in that it only determines with certainty which bands

---

* The author is indebted to the Research Department of Nobel's Explosive Company for this material.

† At ordinary temperatures nitrogen peroxide consists of a mixture of N_2O_4 and NO_2, there being roughly five molecules of the former to every one of the latter. Dissociation into NO_2 is complete at 140° C.

‡ The latter lie at 1333 cm.$^{-1}$ and 814 cm.$^{-1}$, *cf.* Dadieu, Jele and Kohlrausch, SitzBer. Akad. Wiss. Wien., vol. 140, p. 293 (1931).

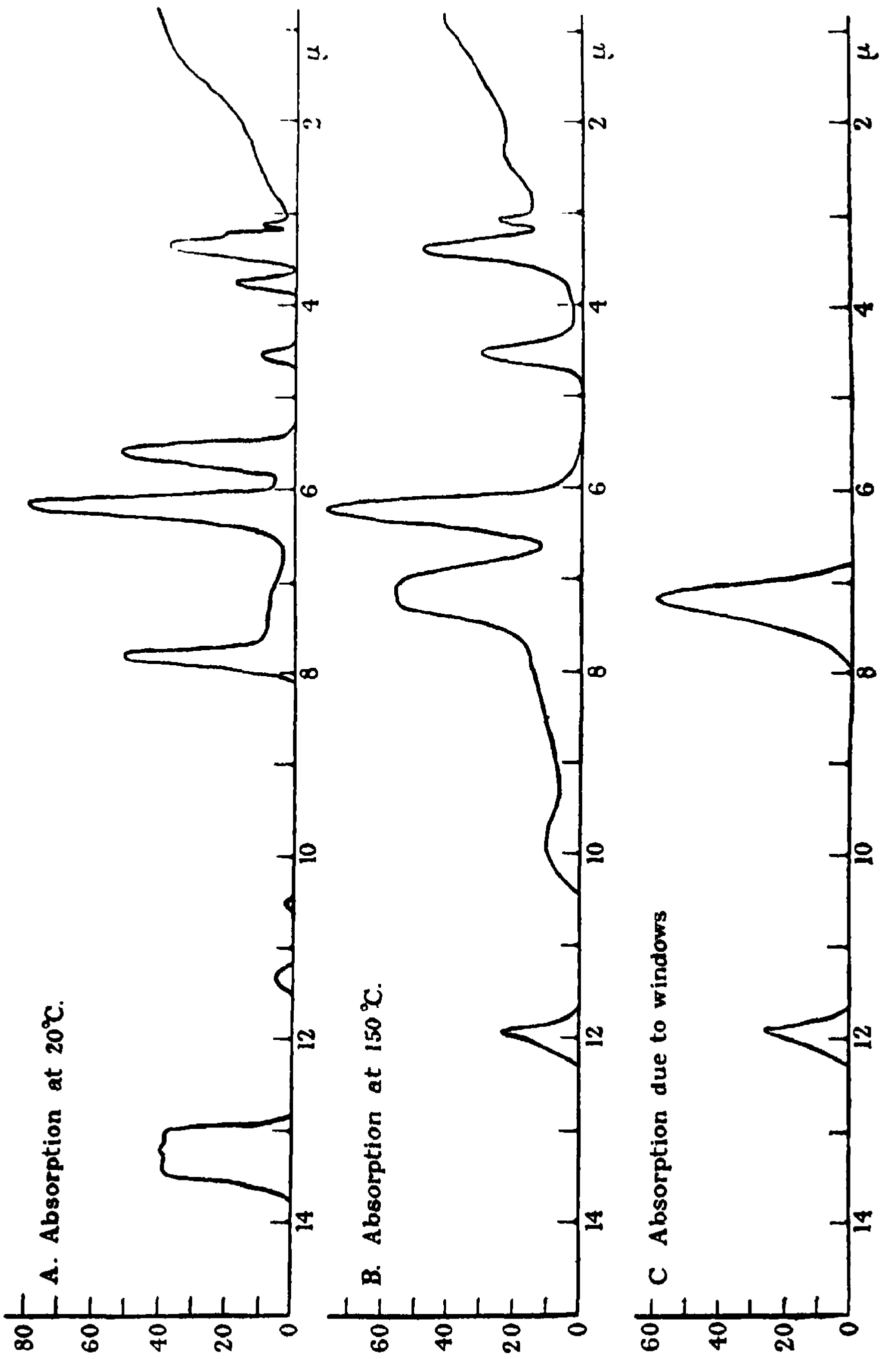

A. Absorption at 20°C.
B. Absorption at 150°C.
C Absorption due to windows
80
60
40
20
0
60
40
20
0
60
40
20
0
14
12
10
8
6
4
2
μ

Fig. 1.

are due to the dioxide. It may be the case that some of the bands of the tetroxide are obscured by the bands of the dioxide, but we shall find when we come to the final interpretation of the combined infra-red and Raman spectra that this is unlikely. We may conclude that the main absorption bands of the two oxides in the region from 3 μ to 14 μ are those given in Table I.

Table I.—Infra-red Absorption Bands of Nitrogen Peroxide.

NO_2.		N_2O_4.	
cm.$^{-1}$	Intensity	cm.$^{-1}$.	Intensity.
1000	Very weak	752	Strong
1350 (?)	Weak	882	Weak
1615	Very strong	948	Very weak
2220	Weak	1265	Strong
2910	Medium	1749	Strong
3120	Medium	2620	Weak
3240	Weak	2975	Medium

The Contours of the Bands.

The contours of the bands were next determined, using grating spectrometers of very high resolving power. The runs were carried out at room temperature, as the preliminary examination had enabled one to decide which bands were characteristic of the different molecules. The amount of gas in the cell varied according to the intensity of the band under examination, being adjusted to give a maximum absorption of about 60 per cent. Only the bands due to the N_2O_4 molecule will be described in this paper.*

The Band at 752 *cm.*$^{-1}$, *fig.* 2.—The band at 13·3 μ was examined first on the prism instrument, with a very small amount of gas in the cell, and definite indications of what may be termed P, Q, and R branches were observed.

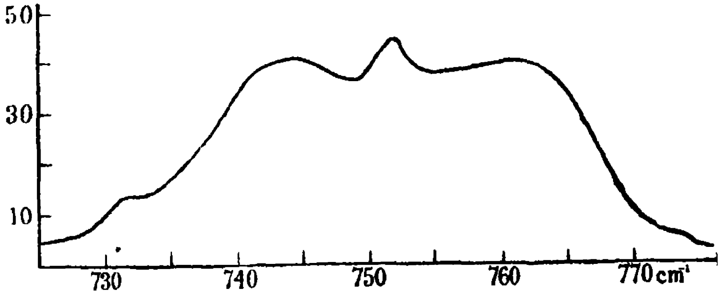


Fig. 2.

---

* A complete description of the dioxide bands and a discussion of the NO_2 molecule is reserved for a later paper.

On re-examination with a grating instrument (using the first order of a grating with 1200 lines to the inch) no sign of resolution into individual rotation lines was observed. The slit width employed corresponded to 0·5 cm.$^{-1}$. The effect was simply to spread out the contour obtained from the prism spectrometer without revealing anything new. The separation between what may be conventionally named the P and R branches was approximately 17 cm.$^{-1}$.

The Band at 1265 *cm.*$^{-1}$, *fig.* 3.—With the prism spectrometer the band at 7·9 μ was exceedingly narrow and intense, but showed no structure whatever; upon examination in the second order of the 1200 line grating definite indications of a "P, Q, R" structure were found, but no resolution into rotation lines. The slit width used corresponded to an interval of 0·6 cm.$^{-1}$; the approximate P-R separation in this band was 14·5 cm $^{-1}$.

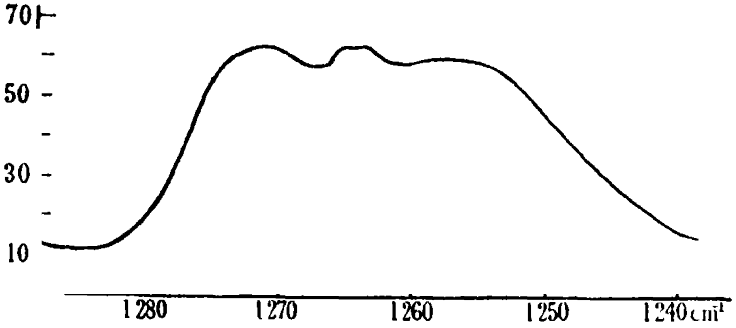


Fig 3

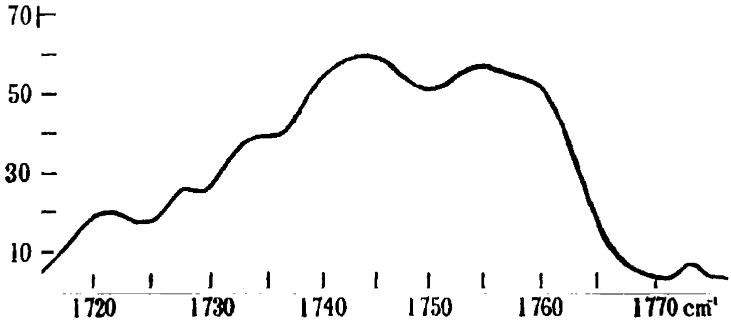


Fig 4.

The Band at 1749 *cm.* $^{-1}$, *fig.* 4.—The band at 5·7 μ is in rather a difficult region, as it overlaps the end of the water vapour band at 6·2 μ. The prism spectrometer revealed it as a very narrow intense band with a single maximum.

Under high dispersion, however, and with small amounts of gas in the cell, a doublet structure was revealed. The definite absence of anything in the nature of a Q branch in this band will be found to have important theoretical consequences. The slit had a width corresponding to an interval of 3·0 cm.$^{-1}$; the approximate doublet separation was 13·5 cm.$^{-1}$.

The Band at 2975 *cm.*$^{-1}$, *fig.* 5.—In the examination of the absorption band at 3·4 μ under high dispersion it was discovered that what had appeared to be a single band under low dispersion was, in fact, two distinct bands. One of these was a simple doublet, due to NO_2, but the other was a narrow unresolved band, which was proved to be due to N_2O_4. The grating used had 7200 lines to the inch, and the slit width was 2 cm.$^{-1}$.

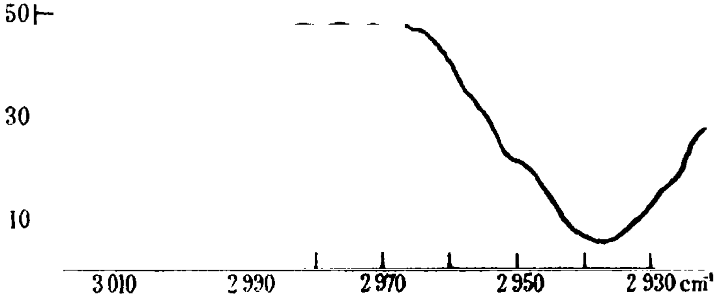


Fig. 5

The Bands at 882 *cm.*$^{-1}$, 948 *cm.*$^{-1}$, *and* 2620 *cm.*$^{-1}$.—These bands have not been examined under high dispersion. They are all very weak and it is quite improbable that any of them is a fundamental frequency of the molecule. A knowledge of their contours would be of little use to us since the theory of the combination bands in this type of molecule is as yet rather undeveloped.

The Structure of the Molecule

From chemical evidence there have been three different theories regarding the structure of this molecule.* These are illustrated by the three diagrams of fig. 6. Each of these forms will give rise to certain characteristic features in the infra-red and Raman spectra. It will be shown that the present data on the infra-red and Raman spectra are overwhelmingly in favour of the structural formula $O_2N{-}NO_2$. From the ease with which the molecules of N_2O_4 dissociate into molecules of NO_2 this formula has for long been preferred to the other two. It has been suggested that the three forms co-exist in an equilibrium mixture, this will be shown to be extremely unlikely.

---

* Mellor, "Treatise on Inorganic Chemistry," vol. 8, p. 546.

Since there are six atoms in the molecule, there are twelve $(3 \times 6 - 6)$ degrees of internal freedom, and therefore twelve fundamental frequencies. We can form a physical picture of these twelve modes of vibration for the O_2N–NO_2 model by considering them to be built up from the mutual vibrations of two NO_2 groups. Each NO_2 group has three fundamental frequencies, so if we consider them to be loosely coupled together six frequencies will result, since each pair can be taken either exactly in phase or exactly out of phase. These frequencies we illustrate in fig. 7 as ν_1, ν_2, ν_3, ν_4, ν_5, ν_6. To obtain the other six frequencies we consider the mutual vibrations of the two NO_2 groups, treated as rigid triangles. This method of building up the normal vibrations by using extreme force fields has been justified by Dennison.* The diagrams in fig. 7 are merely intended to give a rough idea of the type of motion which is to be associated with each frequency. The accompanying table shows which of the frequencies are active in infra-red absorption and which in Raman scattering; the direction of vibration of the electric moment is given for the infra-red frequencies. The selection rules for the infra-red frequencies are derived from the assumption of a symmetrical charge distribution in the molecule, those for the Raman scattering, from an application of Placzek's rules.†

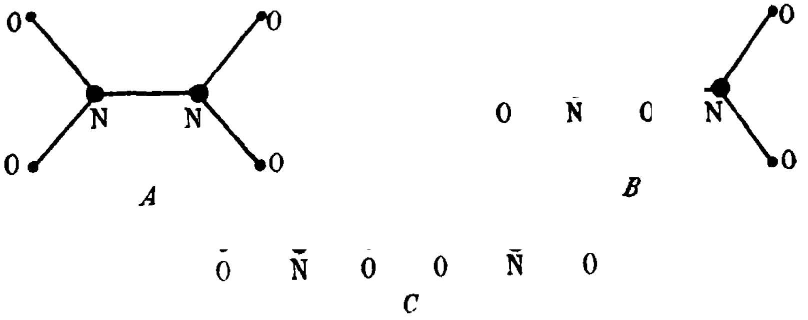


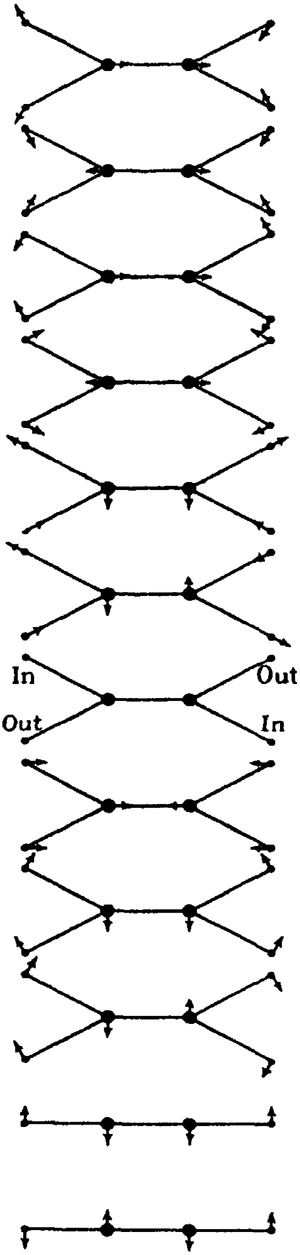
FIG 6

Of the twelve frequencies, only five are active in absorption, and none of these are active in Raman scattering; of the remaining seven, six are active in the Raman effect, but ν_7 is completely inactive. Since the molecule belongs to the class of "asymmetrical top molecules," the character of its fundamental

---

* 'Rev Mod. Phys,' vol. 3, p 280 (1931).

† 'Leipzig Vort,' p. 71 (1931).



	INFRA-RED.	DIRECTION OF ELECTRIC MOMENT.	RAMAN EFFECT.
ν_1	ACTIVE	LEAST AXIS	INACTIVE
ν_2	INACTIVE		ACTIVE
ν_3	ACTIVE	LEAST AXIS	INACTIVE
ν_4	INACTIVE		ACTIVE
ν_5	ACTIVE	MIDDLE AXIS	INACTIVE
ν_6	INACTIVE		ACTIVE
ν_7	INACTIVE		INACTIVE
ν_8	INACTIVE		ACTIVE
ν_9	ACTIVE	MIDDLE AXIS	INACTIVE
ν_{10}	INACTIVE		ACTIVE
ν_{11}	ACTIVE	GREATEST AXIS	INACTIVE
ν_{12}	INACTIVE		ACTIVE

FIG. 7

absorption bands will depend on the direction of vibration of the electric moment in the manner stated by Dennison (*loc. cit.*); and so, of the five fundamental absorption bands, two should possess a well-defined Q branch (vibration along the least axis of inertia), two should possess no Q branch (vibration along the middle axis of inertia), and one should possess a broad ill-defined Q branch (vibration along the greatest axis of inertia). As regards the Raman frequencies, we should expect those corresponding to ν_2, ν_4, ν_8 to be much more intense than those corresponding to ν_6, ν_{10}, ν_{12} since the rate of change of polarizability of the molecule would appear to be much greater for the former than the latter.*

Assignment of the Fundamental Frequencies.

With the above criteria in mind the arrangement of the infra-red and Raman spectra of nitrogen tetroxide in terms of fundamental frequencies and combinations of them is not difficult. The intensity of the infra-red bands at 752 cm.$^{-1}$, 1265 cm.$^{-1}$, and 1749 cm.$^{-1}$ makes the choice of them as fundamental frequencies reasonable and natural. There is little doubt that the other two active infra-red frequencies must lie in the region below 700 cm.$^{-1}$ which has not yet been carefully examined.† Now of the active frequencies, ν_1, ν_3, and ν_5 will lie higher than ν_9 and ν_{11}, being essentially inter-atomic frequencies, whereas the latter are inter-group frequencies, *i.e.*, due to the relative motion of the two NO_2 groups treated as rigid triangles. From the fact that it is the only one with no Q branch we may immediately associate the band at 1749 cm.$^{-1}$ with ν_5. Next we can differentiate between ν_1 and ν_3, for although the contours are similar, each having a well-defined Q branch, ν_1 may be expected to be greater than ν_3 since the group frequency from which it is derived is always greater than that from which ν_3 is derived in the triangular molecules so far investigated. This enables us to fix ν_1 and ν_3 as 1265 cm.$^{-1}$ and 752 cm.$^{-1}$ respectively.

Now consider the Raman spectrum as given in Table II. We shall disregard the frequencies 28 cm.$^{-1}$, 57 cm.$^{-1}$, and 79 cm.$^{-1}$, which are obviously too low to be fundamental vibration frequencies and which are to be discussed else-

---

* *Cf.* Placzek, 'Z. Physik,' vol 70, p. 684 (1931), where it is shown that the intensity of a particular Raman frequency depends on $\partial\alpha/\partial q$ where α is the polarizability and q is the associated normal co-ordinate.

† The very broad band, with its centre at approximately 380 cm.$^{-1}$, observed by Strong and Woo, is probably one of them, although this band has not been proved to be due to N_2O_4 rather than NO_2.

Table II.—Raman Spectrum of Nitrogen Tetroxide* (solid).

Frequency cm $^{-1}$.	28	57	79	283	500	813	1337, 1382, 1724
Intensity	Weak	Weak	Strong	Very strong	Medium	Strong	Medium

where.* We are left with six Raman frequencies, of which the one at 283 cm.$^{-1}$ is much more intense than all the others.† This frequency may immediately be identified with the highly symmetrical vibration ν_8 in which the two halves of the molecule "oscillate against each other." Its character agrees with what might be expected from the fact that the molecule is so easily dissociated. A small binding force between two such heavy groups means that this must be a very low frequency. Again, a small binding force means comparatively large amplitudes in the oscillations and so large changes in the polarizibility, and therefore great intensity in the Raman effect.‡ The other fairly strong Raman line at 813 cm.$^{-1}$ will be either ν_2 or ν_4. Using the same criterion as we employed to differentiate ν_1 from ν_3 we see that the correct choice is ν_4. The absence of a third strong Raman line is a little disturbing at first, but we notice that in the region where it might be expected to fall (as the counterpart of the active frequency at 1265 cm.$^{-1}$), viz., 1300 cm.$^{-1}$, there are two Raman lines of comparable intensity forming a doublet of medium strength. This would appear to be another example of the resonance degeneracy which has been found in other polyatomic molecules.§ What would normally be a single Raman frequency may become a doublet when it happens that the first overtone of a certain other frequency falls very close to it. The molecules in each level instead of possessing the one definite mode of vibration characteristic of that level are divided into two classes, one vibrating in the mode characteristic of the one level, the other vibrating in the mode characteristic of the other level. The proportion of molecules in each class is a function of the separation between ν_A and $2\nu_B$, fig. 8, approaching unity as ν_A becomes exactly equal to $2\nu_B$, and changing rapidly as the separation increases, until when the separation is large, each level is characterized by only one type of vibration. The fact that the two Raman lines at 1337 cm.$^{-1}$ and 1382 cm.$^{-1}$ are so nearly equal in intensity would seem to indicate that the coincidence was very close in this case. We assume, then, an intermediate value of 1360 cm.$^{-1}$ for ν_2. Of the remaining Raman frequencies, the assignment of the one at 1724 cm.$^{-1}$ to ν_6 as the counterpart of the active ν_5 at 1749 cm.$^{-1}$ follows naturally, while the reasons for taking the other one at 500 cm.$^{-1}$ to be ν_{10} rather than ν_{12} will be given in the next section.

---

* Sutherland, 'Proc. Roy Soc.,' A, vol. 141 (1933), in the press.

† Observed by Menzies and Pringle as the only Raman line of nitrogen tetroxide, 'Nature,' vol. 127, p. 707 (1931).

‡ Placzek, 'Z. Physik,' *loc. cit.*

§ Fermi, 'Z. Physik,' vol. 71, p. 250 (1931); Dennison, 'Phys. Rev.,' vol. 41, p. 304 (1932).

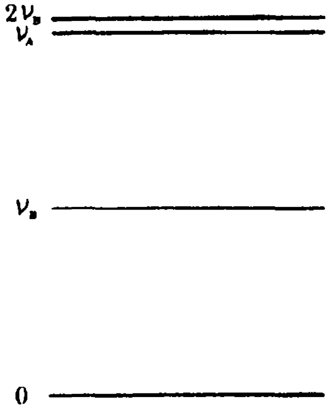


FIG. 8.

These results are summarized below, together with a table showing how all but one of the remaining bands in the infra-red spectrum now receive a satisfactory interpretation. Of course, the values of the Raman frequencies cannot be taken as being correct to more than about 15 cm.$^{-1}$, since they were obtained for the solid, and there is usually a slight increase in the frequency in going from the solid to the gas. Moreover, since there are not sufficient combination bands known to determine the anharmonic and interaction constants, exact agreement between the observed and calculated values of the combination frequencies cannot be expected.

Infra-red fundamentals.

$\nu_1 = 1265$ cm.$^{-1}$
$\nu_3 = 752$ cm.$^{-1}$
$\nu_5 = 1749$ cm.$^{-1}$
ν_9 or $\nu_{11} = 380$ cm.$^{-1}$

Raman fundamentals.

$\nu_2 = 1360$ cm.$^{-1}$
$\nu_4 = 813$ cm.$^{-1}$
$\nu_6 = 1724$ cm.$^{-1}$
$\nu_8 = 283$ cm.$^{-1}$
$\nu_{10} = 500$ cm.$^{-1}$

Combination

Frequency observed.	Frequency assignment.	Frequency calculated.
cm.$^{-1}$		cm.$^{-1}$
882	$\nu_{10} + \nu_9$ or ν_{11}	880
2614	$\nu_1 + \nu_2$	2625
2975	$\nu_1 + \nu_6$	2989

Relations between the Frequencies.

So far the discussion of the spectra of nitrogen tetroxide has been on rather qualitative lines, but if two very simple and very general assumptions are made regarding the forces which hold the molecule together, then certain quantitative relations may be established between the fundamental frequencies. These assumptions are :—

(1) That the potential function for the molecule possesses the same geometric symmetry as does the molecule itself ;

(2) That, to a first approximation, only neighbouring atoms affect one another.

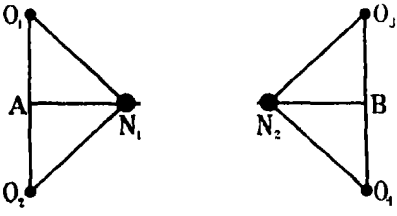


FIG. 9.

Let us choose as displacement co-ordinates for the system—

x_0 the relative displacement of the two N atoms parallel to the line N_1N_2.

x_1 the relative displacement of N_1 and A, the centre of gravity of the two oxygen atoms O_1O_2

x_2 the relative displacement of N_2 and B, the centre of gravity of the two oxygen atoms O_3O_4.

q_1 the relative displacement of O_1 and O_2 perpendicular to N_1N_2.

q_2 the relative displacement of O_3 and O_4 perpendicular to N_1N_2.

y_1 the displacement of N_1 perpendicular to N_1N_2.

y_2 the displacement of N_2 perpendicular to N_1N_2.

ϕ_1 the angle that O_1O_2 makes with the perpendicular to N_1N_2.

ϕ_2 the angle that O_3O_4 makes with the perpendicular to N_1N_2.

These nine co-ordinates all refer to motions in the plane of the molecule. There are three other co-ordinates (corresponding to ν_7, ν_{11}, and ν_{12}) which measure motions out of the plane ; these can be treated independently and will be considered later. We next write down the expressions for the kinetic and potential energies corresponding only to motions in the plane of the molecule. It is clear that the kinetic energy T will be given by an expression of the form:

$$T = \tfrac{1}{2}\{A(x_1^2 + \dot{x}_2^2) + B(q_1^2 + q_2^2) + Cx_0^2 + 2E(x_1x_0 + \dot{x}_2x_0) + 2Fx_1x_2$$
$$+ P(y_1^2 + \dot{y}_2^2) + Q(\phi_1^2 + \phi_2^2) + 2R(y_1\phi_1 + y_2\phi_2) + 2Sy_1y_2 + 2T\phi_1\phi_2\},$$

where A, B, C, E, F, P, and S are functions of the masses of the atoms, while Q, R, and T will also involve y_1 or y_2 or both. The absence of cross-product terms between $x_0x_1x_2q_1q_2$ and $y_1y_2\phi_1\phi_2$ follows from the decomposition of the motions of the atoms into the translational motion of N_1, N_2, A, and B parallel and perpendicular to N_1N_2, and the rotational motion of O_1, O_2 and O_3, O_4 about their respective centres of gravity A and B. The fact that certain terms have the same coefficient follows from symmetry considerations. For the potential energy we start with a general homogeneous quadratic form in the variables, but when we modify it by introducing the assumptions (1) and (2) it will be found to reduce to the expression:

$$V = \tfrac{1}{2}\{a(x_1^2 + x_2^2) + b(q_1^2 + q_2^2) + cx_0^2 + 2d(x_1q_1 + x_2q_2)$$
$$+ p(y_1^2 + y_2^2) + q(\phi_1^2 + \phi_2^2) + 2r(y_1\phi_1 + y_2\phi_2) + 2sy_1y_2\},$$

where a, b, c, . are undetermined constants. The fact that again no cross-product terms occur between $x_0x_1x_2q_1q_2$ and $y_1y_2\phi_1\phi_2$ makes it possible to separate the problem into two parts, one dealing with the frequencies ν_1, ν_2, ν_3, ν_4, ν_8 and involving only $x_0x_1x_2q_1q_2$, the other dealing with the frequencies ν_5, ν_6, ν_9, ν_{10} and involving only $y_1y_2\phi_1\phi_2$.

Considering first the frequencies which concern motion along the axis of the molecule we have,

$$T = \tfrac{1}{2}\left\{\frac{2m(M+m)}{2m+M}(x_1^2 + x_2^2) + \frac{2m+M}{2}x_0^2 + \frac{m}{2}(q_1^2 + \dot{q}_2^2)\right.$$
$$\left. + 2m(\dot{x}_1\dot{x}_0 + x_2x_0) + \frac{4m^2}{2m+M}x_1x_2\right\},$$

where M is the mass of an N atom, and m is the mass of an O atom while

$$V = \tfrac{1}{2}\{a(x_1^2 + x_2^2) + b(q_1^2 + q_2^2) + cx_0^2 + 2d(x_1q_1 + x_2q_2)\}.$$

Since there are five frequencies involved, but only four independent constants in the potential energy expression, it follows that one relation must exist between the five frequencies. Proceeding in the usual way we set up the determinantal equation of the normal vibration problem. From its solution we may readily deduce the following relation* between the frequencies, viz.,

$$\frac{\mathrm{M}+2m}{\mathrm{M}}\,\frac{\nu_8{}^2\nu_2{}^2\nu_4{}^2}{\nu_1{}^2\nu_3{}^2}=\nu_8{}^2+\nu_2{}^2+\nu_4{}^2-\nu_1{}^2-\nu_3{}^2.$$

Substituting the values assigned to these frequencies we find that the agreement between the two sides of the equation is surprisingly good considering the small uncertainties in the values of the frequencies.† This must be considered a very satisfactory confirmation of the foregoing work.

When we consider motions perpendicular to the axis of the molecule (but still in the plane of the molecule), we notice that there are four arbitrary constants in the potential energy expression, and since just four frequencies are determined by these it is not possible to find any relations between them.

The problem of treating the motions out of the plane of the molecule may be simplified in a similar manner to that just shown by separation into a twisting motion of the two NO_2 groups about the axis, defined by a co-ordinate ψ, say, and a motion of the atoms perpendicular to the axis of the molecule, corresponding to and illustrated in the diagrams of the frequencies ν_{11} and ν_{12}. The latter two frequencies are quite analogous to the two perpendicular vibrations in acetylene. Methods of treating these two frequencies for acetylene have been given by Kramers and Olson,‡ and by Mecke (*loc. cit.*), and need not be

---

* A similar relation has been shown to hold between the corresponding five frequencies in ethylene by Mecke (' Z. phys. Chem.,' B, vol. 17, p. 1 (1932)). The proof given above is, we feel, more general than that given by Mecke in that it uses less specific assumptions regarding the forces in the molecule.

† The actual values obtained were

$$\frac{\mathrm{M}+2m}{\mathrm{M}}\,\frac{\nu_8{}^2\nu_2{}^2\nu_4{}^2}{\nu_1{}^2\nu_3{}^2}\bumpeq 360{,}000$$

$$\nu_8{}^2+\nu_2{}^2+\nu_4{}^2-\nu_1{}^2-\nu_3{}^2\bumpeq 430{,}000$$

a difference of less than 17 per cent. The sensitivity of the equation to the values of the frequencies is shown by the fact that if we choose ν_2 to be 1337 cm.$^{-1}$ this gives a difference of less than 6 per cent., while if we try the pair of frequencies at 1750 cm^{-1} as ν_1 and ν_2 we obtain

$$\frac{\mathrm{M}+2m}{\mathrm{M}}\,\frac{\nu_8{}^2\nu_2{}^2\nu_4{}^2}{\nu_1{}^2\nu_3{}^2}\bumpeq 300{,}000$$

and

$$\nu_8{}^2+\nu_2{}^2+\nu_4{}^2-\nu_1{}^2-\nu_3{}^2\bumpeq 75{,}000$$

‡ ' J. Amer. Chem. Soc.,' vol. 54, p. 136 (1932).

repeated here. We shall only quote the result which is arrived at by all and is that

$$\frac{\delta_1}{\delta_2} = \left(\frac{m_1 + m_2}{m_1 j^1 + m_2}\right)^{\frac{1}{2}},$$

where

δ_1 is the active frequency,

δ_2 is the inactive frequency,

m_1 is the mass of each of the end atoms,

m_2 is the mass of each of the centre atoms, and

$$j = \frac{l + d}{l}.$$

$2l$ being the distance between the centre atoms,

d being the distance from a centre atom to an end atom.

We may apply this formula to the molecule of nitrogen tetroxide and obtain

$$\frac{\nu_{11}}{\nu_{12}} = \frac{23}{16j^2 + 7}$$

We use as a value for $2l$ that found from the Raman spectrum of nitrogen,* viz., 1·089 A.; for d we take the value for the internuclear distance found from the electronic spectra of nitric oxide,† viz., 1·16 A. multiplied by a factor $1/\sqrt{2}$ (taking the angle between the NO bands as 90°) and obtain as an approximate result that

$$\frac{\nu_{11}}{\nu_{12}} = 0 \cdot 462.$$

It follows immediately that if the Raman frequency at 500 cm.$^{-1}$ is chosen as ν_{12} then there ought to be a strong infra-red band near 230 cm.$^{-1}$, corresponding to the active fundamental ν_{11}. The curves of Strong and Woo (*loc. cit.*) indicate that there is no absorption whatever in this neighbourhood. An additional argument against taking $\nu_{12} = 500$ cm.$^{-1}$ is that the corresponding Raman frequency is too weak to appear in acetylene (and in all probability, also in ethylene). On the other hand if we take $\nu_{10} = 500$ cm.$^{-1}$ then since ν_9 may be expected from general considerations to lie slightly lower, it seems reasonable to interpret the band observed by Strong and Woo (*loc. cit.*) near 380 cm.$^{-1}$ as ν_9. If finally we assume‡ that ν_{11} is the fundamental of which the overtone coincides with the frequency ν_2 at 1360 cm.$^{-1}$, *i.e.*, that $\nu_{11} = 680$ cm.$^{-1}$, then ν_{12} may be calculated to lie at 1480 cm.$^{-1}$. In view of the fact that this relation between ν_{11} and ν_{12} is probably only very approximately correct* this value for ν_{12} must be accepted with caution.

---

* Rasetti, 'Nature,' vol. 123, p. 757 (1929).

† Jenkins, Barton, and Mulliken, 'Phys. Rev.,' vol. 30, p. 150 (1927).

‡ This assumption is given considerable support from the fact that in a preliminary survey of the infra-red spectrum of nitrogen peroxide between 14μ and 20μ by Mr. Tayou Wu a new absorption band was picked up with its centre at 682 cm.$^{-1}$ (private communication). Additional support is given by the fact that the very weak band at 948 cm.$^{-1}$ now receives an interpretation as $\nu_8 + \nu_{11}$.

We have given here one possible assignment for the values of ν_9, ν_{10}, ν_{11}, and ν_{12}, and although on the present evidence it appears to us the most satisfactory one, we cannot be certain that the alternative interpretation with $\nu_9 = 680$ cm.$^{-1}$, $\nu_{11} = 380$ cm.$^{-1}$, and $\nu_{12} = 500$ cm.$^{-1}$ is not the correct one In the latter case we have no means of estimating the value of ν_{10} since no numerical relation exists between it and any of the other frequencies. As the inactive counterpart of the active frequency ν_9 it probably lies a little higher than ν_9 but its value may be anywhere between 800 cm.$^{-1}$ and 1000 cm.$^{-1}$.

The determination of the remaining fundamental ν_7 from the existing spectral data would appear to be impossible. Its magnitude cannot be related to any of the other eleven frequencies since it depends on only one force constant of which these are all independent. It can only appear as a combination band, and the observed combination bands can be readily explained in terms of the frequencies already determined. The only clue to its value will be from specific heat data which we now proceed to consider

Specific Heat of Nitrogen Tetroxide.

The calculation of the specific heat of nitrogen tetroxide is of considerable theoretical interest in view of the large contributions to the vibrational part of it which are to be expected from the numerous low fundamental frequencies in this molecule. It is important also in that an accurate value of the specific heat of nitrogen tetroxide is very difficult to obtain experimentally and yet is often required in physical and chemical investigations on the molecule. It is well known that the specific heat (at constant volume) of any molecule is made up of three parts corresponding to the three different types of energy which a molecule may possess, viz., translational vibrational, and rotational energy. Thus we write

$$C_v = C_T + C_{\text{vib.}} + C_{\text{rot.}}$$

Now at room temperatures the distribution of the molecules in the various translational and rotational energy states is so nearly classical that we may take $C_T + C_{rot} = 3R$, but C_{vib} will be made up of a number of terms of which a typical one is

---

* The corresponding relation is by no means well satisfied for acetylene.

$$C_{vib}\,(\nu) = R\left(\frac{\beta\nu}{T}\right)^2 \frac{e^{\frac{\beta\nu}{T}}}{\left(e^{\frac{\beta\nu}{T}} - 1\right)^2},$$

where ν is any fundamental frequency of the molecule $\beta = h/\kappa = 4 \cdot 77 \times 10^{-11}$, and T is the absolute temperature.

When ν is above 1000 cm.$^{-1}$ the value of C_{vib} (ν) at ordinary temperatures is practically negligible, but as ν decreases the contribution increases very rapidly, so that for $\nu = 500$ cm.$^{-1}$ it is 1·27 cals. per mole, while for $\nu = 200$ cm^{-1} it is approximately 1·83 cals per mole. This means that the value of the specific heat may be quite sensitive to changes in the values of the fundamental frequencies lying between 500 cm.$^{-1}$ and 1000 cm.$^{-1}$ and each fundamental below 850 cm.$^{-1}$ will add at least 0·5 cal. per mole, so that an accurate experimental value of the specific heat may serve as a useful discriminant between different frequency assignments.

The value of the vibrational specific heat as calculated from the fundamental frequencies given in Table III is 6·8 cals. per mole at a temperature of 300° A. If the alternative assignment of fundamentals is taken with $\nu_9 = 680$ cm^{-1}, $\nu_{11} = 380$ cm.$^{-1}$, and $\nu_{12} = 500$ cm.$^{-1}$, we arrive at the value of 7·5 cals. per mole. The total specific heat is then 12·8 cals. per mole in the first case and

Table III.—Fundamental Frequencies of Nitrogen Tetroxide.

$\nu_1 = 1265$ cm^{-1}	$\nu_4 = 813$ cm^{-1}		$\nu_{10} = 500$ cm.$^{-1}$
$\nu_2 = 1360$ cm^{-1}	$\nu_5 = 1749$ cm^{-1}	283 cm^{-1}	$\nu_{11} = 680$ cm.$^{-1}$
752 cm.$^{-1}$	$\nu_6 = 1724$ cm^{-1}	380 cm^{-1}	$\nu_{14} = 1480$ cm^{-1}

13·5 cals. per mole in the second If the experimental value of the specific heat were sufficiently accurate we ought to be able to obtain at least a rough estimate of ν_7, supposing that it is less than 800 cm.$^{-1}$. The probability of ν_7 being greater than 800 cm.$^{-1}$ can be taken as rather small when we consider the magnitudes of the other inter-group frequencies.

Unfortunately the experimental data on the specific heat of nitrogen tetroxide are most unreliable. The values collected by Partington and Shilling* (based on measurements of the velocity of sound) are vitiated by the unknown contribution which comes from the heat of dissociation of the molecule. Recently

---

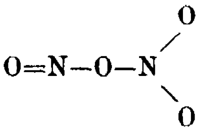
* "The Specific Heats of Gases" (1924), p. 188.

McCollum* has measured the specific heat by a flow method and tried to estimate the contribution due to dissociation. The latter is so large, however, that errors of 1 or 2 per cent. in it may lead to values of the specific heat differing by 20 per cent. or more. From the table of re-calculated values in Partington and Shilling's book an average value for C_v at 300° A., and with 20 per cent. dissociation into NO_2, is 19 cals. per mole. Taking C_v for NO_2 to be 7 cals. per mole the value of C_v for N_2O_4 may be obtained from the equation

$$C_v \text{ (apparent)} = \tfrac{4}{5} C_v (N_2O_4) + \tfrac{2}{5} C_v (NO_2)$$

as approximately 20 cals per mole. Such a value is too high, because, as we remarked, no account has been taken of the heat of dissociation. On the other hand, the value obtained by McCollum is certainly too low. He gave for C_p at 306·7° A., with 24 per cent. dissociation, the value 11·4 cals. per mole. This corresponds to a value for $C_p (N_2O_4)$ close to 9·5 cals. per mole and so to an approximate value of 7·5 cals. per mole for $C_v (N_2O_4)$. Such a value is obviously too low in a molecule which is definitely known to possess fundamentals as low as 283 cm.$^{-1}$, 500 cm.$^{-1}$, and 752 cm.$^{-1}$ As a result of such uncertainty in the experimental results no estimate of the value of ν_7 can be made. A careful experimental redetermination of the specific heat of nitrogen tetroxide would be of great interest in the light of the above discussion. Meanwhile we can give a theoretical figure for the specific heat of nitrogen tetroxide, viz., 14 cals. per mole which is possibly correct to within 10 per cent.

Discussion of other Possible Structures.

The correlation of the spectra of nitrogen tetroxide in terms of the fundamental frequencies of the model A of fig. 6 is so satisfactory that the consideration of other possible structures might seem tedious and unnecessary. It may be shown very briefly, however, that neither of the other structures illustrated in fig. 6 will explain even the outstanding features of the infra-red and Raman spectra. Consider first the linear structure O=N—O—O—N=O; such a molecule would undoubtedly exhibit a strong Raman spectrum as a result of its high symmetry. If only the Raman spectrum were known then it would not be possible to choose between this structure and the first one. But when we remark that the three highest, active, fundamental frequencies of this type of molecule would be parallel frequencies (*i.e.*, vibration of the electric moment parallel to the axis of the molecule), and therefore would exhibit a simple doublet structure with identical spacings, whereas we observe in the three highest fundamentals that two possess Q branches and that the doublet spacing is by no means constant, we may conclude that such a structure is definitely ruled out of consideration. A molecule with the third structure proposed,  $\mathrm{O{=}N{-}O{-}N}\begin{array}{l}\diagup \mathrm{O}\\ \diagdown \mathrm{O}\end{array}$ being rather unsymmetrical would not be expected to have a strong Raman spectrum*; the existence of the very strong line at 283 cm.$^{-1}$ would be especially hard to justify. The same lack of symmetry would give rise to a very rich spectrum in the infra-red, for instead of only five active frequencies we would expect all twelve fundamentals to appear in the infra-red. The comparative scarcity of strong absorption bands in the infra-red is by itself sufficient evidence against this structure. The theory that nitrogen tetroxide exists as an equilibrium mixture of the three forms is now clearly untenable.

---

* 'J. Amer. Chem. Soc.,' vol. 49, p. 28 (1927).

There is still one possibility which we have neglected to consider; that is, that the molecule is not plane but that the two NO_2 groups are in planes at right angles to each other. Such a molecule will belong to the "symmetrical top" class of molecules since it has two of its principal moments of inertia equal. Its infra-red bands will be divided into two classes according as the change of electric moment takes place along, or perpendicular to the symmetry axis. The former resemble the bands of a diatomic molecule, having rotation lines equally spaced by an amount $h/4\pi^2C$ (C being the moment of inertia perpendicular to symmetry axis), but in addition possess a strong Q branch. The structure of the latter is more complex, and the resulting contour† will depend on the ratio of the moments A and C. In such a molecule there will only be 10 fundamental frequencies (two being doubly degenerate) since the frequencies ν_9 and ν_{11}, and similarly ν_{10} and ν_{12} become identical. Thus we should have only four fundamental infra-red absorption bands and none of these would exhibit a simple doublet structure such as is found in the band at 1750 cm.$^{-1}$. An application of Placzek's rules shows that the Raman spectrum of such a molecule would not be so strong but would contain more lines than that of the plane molecule. The evidence against the staggered form of molecule would thus seem to be fairly conclusive.

---

* *Cf.* Formaldehyde and other molecules of this type in the summary by Kohlrausch "Smekal-Raman-Effekt" (1931).

† Dennison and Gerhard, 'Phys. Rev.,' vol. 43, p. 197 (1933).

The Constants of the Molecule.

One of the definite pieces of information we have come to expect from an investigation of the infra-red spectrum of a molecule is an accurate estimate of its moments of inertia. Strictly speaking, this is possible only for linear molecules, and although for "symmetrical top" molecules the value of one of the moments can be found, for "asymmetrical top" molecules no definite relation has yet been deduced between the spacing of the intensity maxima and the moments of inertia of the molecule. This being so, we prefer not to attempt what would be at best rather unsatisfactory estimates.

It is possible, however, to evaluate some of the constants in the potential energy expression which we have assumed for the molecule, and of particular interest is the coefficient c which relates to the binding force between the two NO_2 groups. The value of c as calculated from the solution of the determinantal equation is $1 \cdot 5 \times 10^5$ dynes per centimetre. Such an abnormally low value* is consistent with the ease of dissociation and with the observed low heat of dissociation,† viz., 13 k. cals. per mole. It is worth noting that when the latter is expressed in wave numbers (4580 cm.^{-1}) it gives a value not too improbable as the convergence limit for overtones of the fundamental frequency ν_8 at 283 cm.^{-1}.

Acknowledgments.

It is a pleasure to record the thanks due to Professor T. M. Lowry under whose direction the problem was begun, and to Dr C. P. Snow for many helpful discussions at an early stage The main portion of the work was done during the tenure of a Commonwealth Fund Fellowship at the University of Michigan, U.S.A., where the author had all the facilities of an excellent laboratory most generously put at his disposal, and where the continued interest and advice of Professor D. M. Dennison have been invaluable.

Summary.

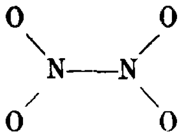
(1) An account is given of new experimental work on the infra-red spectrum of nitrogen peroxide It has been possible to distinguish the absorption bands due to the N_2O_4 molecule from those due to the NO_2 molecule, and the main bands due to the former have been examined under high resolution.

---

* *Cf.* the corresponding c in N_2 which is $22 \cdot 2 \times 10^5$ dynes per centimetre.

† Bodenstein, 'Z. phys. Chem.,' vol. 100, p. 68 (1922).

(2) The normal vibrations of a molecule of the form



are discussed and the main features of its infra-red and Raman spectra are predicted.

(3) The new data on the infra-red and Raman spectra of nitrogen tetroxide are critically compared with these predictions and found to receive a very satisfactory interpretation.

(4) Certain relations between the twelve fundamental frequencies are deduced by making two very simple assumptions regarding the nature of the force fields in the molecule. These relations are used to confirm the assignments based on more qualitative arguments and enable one to determine eleven of the twelve fundamental frequencies.

(5) A discussion of the specific heat of nitrogen tetroxide is given as a means of determining the twelfth fundamental. On account of the uncertainty in the experimental data it is impossible to obtain even a rough estimate of this frequency. A theoretical figure is given for the specific heat, viz., 14 cals. per mole which is possibly correct to 10 per cent

(6) Other models are discussed for the molecule and it is shown that these all fail to account for even the main features of the observed spectra.

(7) The binding constant between the two NO_2 groups is calculated to be $1 \cdot 5 \times 10^5$ dynes per centimetre; such a low value is consistent with the observed dissociation properties of the molecule.

A New Presentation and Interpretation of the Quantum Equations.

By H. T. Flint, Reader in Physics in the University of London, King's College.

(Communicated by O W. Richardson, F R.S.—Received February 18, 1933)

Introduction.

The attempt to unite the first order equations of the quantum theory with the theory of relativity has led to a number of suggestions by various writers, most of whom consider operations in a continuum with four co-ordinates (x, y, z, t) Some add a fifth co-ordinate in order to complete the description of a point. For example, Fock and Iwanenko* add ζ, a quantity capable of possessing the integral values (1, 2, 3, 4).

Einstein† in his latest attempt at a unitary theory retains the four dimensional continuum as the fundamental background, but introduces vectors with five components which are related to four dimensional vectors by means of a set of quantities γ, with appropriate affixes. This is a position half way between the four and five dimensional continuum and, in making use of this system, Solomon‡ has found it necessary to introduce a fifth co-ordinate in order to include quantum phenomena into the scheme.

The view taken here and in earlier papers is that five co-ordinates give the most satisfactory system for the description of the phenomena studied in physics, especially is this so in the development of a unitary theory.

We adopt, therefore, the form of the continuum proposed by Kaluza and Klein,§ with the modification proposed by Fisher.||

It is well known that the track of a photon in the theory of relativity is a null geodesic, and that, associated with the equation of the geodesic, we have a wave equation, which is the equation of a light wave. The first order equations, from which this wave equation is deduced, are Maxwell's electromagnetic equations for empty space.

In the continuum we employ here the null geodesics are the tracks of protons, electrons and photons.

---

* 'Z. Physik,' vol. 54, p 798 (1929).

† 'SitzBer. Preuss. Akad Wiss ,' vol. 25, p. 541 (1931).

‡ 'J. Phys. Radium,' vol. 3, p. 455 (1932).

§ 'Z. Physik,' vol. 46, p. 188 (1928).

|| 'Proc. Roy. Soc.,' A, vol. 123, p. 490 (1929).

Associated with these tracks is a wave equation, which is the wave equation of the quantum theory, and the set of first order equations from which this is deduced is the set of first order equations of the quantum theory. In this notation it appears that the null geodesic is the basis of the wave theory of matter.

In order to develop this form of the theory it was necessary to introduce a system of metrics into the five dimensional Riemannian system. When this is introduced, after the manner of the proposals of Weyl and Eddington, the first order equations are in the form of a vanishing divergence.*

Weyl and Eddington introduced a system of metrics in order to include the electromagnetic theory into the theory of relativity. This led to the introduction of the function ϕ (with components ϕ_μ) which could be identified as the electromagnetic potential.

In the present case ϕ_μ is replaced by the generalized momentum $\Pi_\mu = p_\mu + e\phi_\mu$, where p_μ is the ordinary momentum of the particle, to which the charge e is attached. Π_μ in the five dimensional continuum is the generalization of p_μ in four dimensions.

The fifth co-ordinate, when it occurs, does so very simply, in all the quantities considered, and the operation $\partial/\partial x^5$ is equivalent to multiplication by $2\pi i m_0 c/h$. This suggests a periodicity associated with this co-ordinate and further study indicates the discontinuous occurrence of this variable, which appears in multiples of h/m_0c, m_0 denoting the rest-mass of the particle.

This discontinuous introduction of x^5 reminds us of Fock and Iwanenko's ζ, although there is a difference in the method adopted by these writers.

The discontinuity of x^5 is connected with a discontinuity along the null geodesic and leads to a principle of minimum proper time and to a principle of indeterminateness.†

This principle of indeterminateness, which refers to a definite indeterminacy in position, was also discovered by Ruark‡ and has been established in a variety of ways since by many writers.

The discovery that Dirac's equations were able to be referred to equations which expressed the vanishing of a divergence was an important step towards the object we had in view, but we could only attain that object by showing that the quantity concerned in the equation had some geometrical significance in the continuum.

---

* Flint, 'Proc. Roy. Soc.,' A, vol. 126, p. 644 (1930).

† Flint, 'Proc. Roy. Soc.,' A, vol. 117, pp. 630, 637 (1928).

‡ 'Phys. Rev.,' vol. 31, p. 344 (1928).

This has only been demonstrated indirectly by developing Maxwell's equations from the vanishing divergence.*

The tensor, A, whose divergence vanishes, has been related to the electromagnetic potential and, the latter being a geometric quantity in the continuum, it follows that A is a similar quantity.

Finally the assymetry of the masses of the proton and electron was referred to a metrical property of the continuum.†

The unity thus introduced seems to justify the hope that the interrelation of the fundamental constants of physics, *e g.*, the relation between the gravitational constant and e, h, m_0, and c, will ultimately emerge.

The vanishing divergence referred to above gives first order equations rather more general than those of Dirac, but a slight restriction upon the components of the tensor, A, leads directly to those equations. If this restriction be not made we find that the radius of curvature of the universe enters into the equations, and the result suggests that the restriction should not be too hastily made.‡

It is interesting to note that recent work of Schroedinger§ leads to a more general form of Dirac's equation, which also introduces this radius of curvature.

There is a great difference in the two methods of arriving at these results, and, in comparing the methods, we are able to throw some light on the nature of the quantum problem. It appears worth while to do so now that some evidence is forthcoming that equations containing the radius of curvature can be used to deduce results not before obtainable from the quantum theory.|| It may thus be possible to decide between the various possible notations proposed and to obtain experimental evidence for or against the hypotheses upon which they rest.

An Alternative Form for the Fundamental Matrices.

We begin with some ideas which lead from the matrix notation to ours.

Matrix operations are similar to those occurring between vectors associated with Euclidean space.

Similar operations between tensors are characterized by the introduction of gravitational components.

---

* Fahmy, 'Proc. Phys. Soc.,' vol. 44, p. 368 (1932).

† Flint, 'Proc. Roy. Soc.,' A, vol. 131, p. 170 (1931).

‡ Fahmy, 'Proc. Phys. Soc.,' vol. 45, p. 67 (1933).

§ 'SitzBer. Preuss. Akad. Wiss.,' p. 105 (1932).

|| McVittie, 'Mon. Not. R. Astr. Soc.,' vol. 92, p. 868 (1932)

Thus, if we denote by ε_a a unit vector in Euclidean space, we have

$$\varepsilon_a \cdot \varepsilon_b = \delta_b^a$$

(*i.e.*, $= 0$ or 1 according as $a \neq b$ or $a = b$).

If i_μ denote the fundamental vector in a curvilinear system, $i_\mu \,.\, i_\nu = g^{\mu\nu}$, where $g^{\mu\nu}$ is a gravitational component.

A fundamental operator in the simpler vector analysis is $\varepsilon_a \varepsilon_a$, summation over a being implied.

This is the unit operator converting a vector into itself. A more general operator is $\Sigma\Sigma\, \varepsilon_a \varepsilon_b$, the summation being over a and b.

Let us introduce here a generalization of the Kronecker delta, δ_b^a, just introduced.

These quantities have the values 0, + 1 or − 1. They have, in the cases we require to consider, two superscripts and two subscripts. If the former are distinct from one another and the latter the same two numbers as the superscripts, the value is + 1 or − 1 according as the two numbers are in the same order or not. In all other cases the value is zero.

Thus $\delta_{34}^{12} = 0$, $\delta_{12}^{12} = 1$, $\delta_{21}^{12} = -1$.

We shall also introduce the quantity, d_{mn}^{ab}, with the property that the value is unity provided a, b and m, n are the same numbers, independently of the order. In other cases the value is zero.

By means of these operators we can readily pick out terms from the general operator, $\Sigma\,\Sigma\, \varepsilon_a \varepsilon_b$, and can vary the sign of the coefficients.

Thus the matrix

$$\sigma_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

can be replaced in our notation by

$$\sigma_1 = d_{ab}^{12} \varepsilon_a \varepsilon_b + d_{ab}^{34} \varepsilon_a \varepsilon_b,$$

and similarly

$$\sigma_2 = \begin{matrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{matrix}$$

can be replaced by

$$\sigma_2 = - i\, \delta_{ab}{}^{12}\varepsilon_a\varepsilon_b - i\, \delta_{ab}{}^{34}\varepsilon_a\varepsilon_b.$$

The notation implies, as usual, summation over the possible values of a and b, which, in the first term, are 1, 2 for each, and in the second 3, 4 for each.

Operations between such quantities depend on scalar multiplications between the fundamental vectors $\varepsilon_a\varepsilon_b$.

We can illustrate this by introducing the matrix

$$\rho_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

or

$$\rho_1 = d_{ab}{}^{31}\varepsilon_a\varepsilon_b + d_{ab}{}^{24}\varepsilon_a\varepsilon_b,$$

and considering the scalar operation ·

$$\begin{aligned} \alpha_1 &= \rho_1\sigma_1 \\ &= (d_{ab}{}^{31}\varepsilon_a\varepsilon_b + d_{ab}{}^{24}\varepsilon_a\varepsilon_b) \,.\, (d_{mn}{}^{12}\varepsilon_m\varepsilon_n + d_{mn}{}^{34}\varepsilon_m\varepsilon_n) \\ &= d_{ab}{}^{23}\varepsilon_a\varepsilon_b + d_{ab}{}^{14}\varepsilon_a\varepsilon_b, \end{aligned}$$

which corresponds to the matrix ·

$$\alpha_1 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

The notation has been adopted to show the relations of these vector operators to Dirac's matrices.

It thus appears that the components of the matrices can be obtained from the operators :

$$\pm \Sigma\Sigma\, \varepsilon_a\varepsilon_b, \quad \pm i\Sigma\Sigma\, \varepsilon_a\varepsilon_b,$$

and we can imagine a method of treatment in which these operators play the fundamental rôle instead of groups of operators such as α_1, which are composed of some terms from these more general operators.

The Generalization of the Operators.

Attempts have been made to generalize Dirac's theory in order to bring it into union with the general theory of relativity.

In these attempts we have, side by side, matrix operations, which may be compared with operations between the orthogonal vectors ε_a, and operations introducing gravitational components, *i.e.*, tensor operations.

This occurrence of the two types of operator suggests the occurrence of orthogonal vectors within a curvilinear system such as Einstein* introduced in his theory of parallelism.

In this system a vector may be expressed in two ways: by reference to the orthogonal system and by reference to the curvilinear system.

Each orthogonal unit vector ε_a has components along the curvilinear vectors, which we can write $h^{\mu a}$ so that

$$\varepsilon_a = h^{\mu a} i^{\mu},$$

or in terms of covariant components

$$\varepsilon_a = h_{\mu a} i_{\mu}$$

Any vector, A, may be expressed as $A_a \varepsilon_a$ or $A^{\mu} i^{\mu}$ and to pass from the former to the latter we apply the operator $h^{\mu a} i^{\mu} \varepsilon_a$ as follows:

$$A^{\mu} i^{\mu} = h^{\mu a} i^{\mu} \varepsilon_a \,.\, A_b \varepsilon_b = h^{\mu a} A_a i^{\mu},$$

or

$$A^{\mu} = h^{\mu a} A_a,$$

in accordance with the results of Einstein.

This illustrates the property of the new operator and it occupies a fundamental position like the unit operators of the ordinary vector calculus.

We may thus expect a replacement of the operator $\Sigma\,\Sigma\, \varepsilon_a \varepsilon_b$ by one in which the *h*'s occur.

This is foreshadowed in Fock's paper† in which his matrix α_k is replaced by $e_k h^{\sigma k} \alpha_k$, where e_k has the value ± 1 according to the value of the integer *k*.

We shall regard the operator $h^{\mu a} i^{\mu} \varepsilon_a$ as replacing these matrices, which are, in fact, obtained by taking some of its components.

Schroedinger has recently developed a matrix notation by the introduction of matrices γ_{μ} which appear as fundamental quantities. He states that he wishes to avoid the use of the "less trusted and decidedly more inconvenient form" of the orthogonal components of Einstein. It is his purpose to treat the matrices as fundamental

---

* 'SitzBer. Preuss. Akad. Wiss.,' pp. 217, 224 (1928).

† 'Z. Physik,' vol. 57, p. 266, equation (23) (1929).

One must agree with Schroedinger that the combination of orthogonal components and matrices is inconvenient in its form. We avoid this here by regarding the operator as fundamental instead of parts of it. He writes also that it is not very clear whether the Einstein idea of parallelism plays any part or whether the methods adopted are independent of it

Fock appears to use the Riemannian variation in his paper and then discusses the mode of variation of a quantity described as a "half-vector." He obtains results very similar to those of Schroedinger.

The present paper is a continuation of an earlier paper* in which it was shown that the second order equation of the quantum theory could easily be expressed in Einstein's notation and that quantum mechanics indicated a special choice of orthogonal components. In the first paper as well as in the present one we definitely adopt Einstein's theory of parallelism, but we make use of a system of five co-ordinates. It is thought that in this way we unite both the theories recently proposed by Einstein and, in so doing, make possible the inclusion of physical phenomena into a single theory.

Schroedinger's matrices γ_μ or γ^μ correspond to our quantities $h_{\mu a}\varepsilon_a$ or $h^{\mu a}\varepsilon_a$, which are no more than the expression of the curvilinear vectors in orthogonal co-ordinates.

We do not imply that γ_μ is equal to $h_{\mu a}\varepsilon_a$, but the two quantities play similar parts in the two notations. A property of the h's is

$$h_{\mu a}h_{\nu a} = g_{\mu\nu}. \tag{1}$$

This corresponds to Tetrode's† relation which Schroedinger uses

$$\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2g_{\mu\nu}. \tag{2}$$

The operator, $\gamma_\mu\gamma_\nu$, appears here as similar to $h_{\mu a}h_{\nu b}\varepsilon_a\varepsilon_b$, so that the form (2) gives an operator symmetrical in μ and ν, viz.,

$$(h_{\mu a}h_{\nu b} + h_{\nu a}h_{\mu b})\,\varepsilon_a\varepsilon_b, \tag{3}$$

and the spur of this, *i.e.*, the quantity obtained by multiplying the vectors ε_a and ε_b scalarly, is

$$h_{\mu a}h_{\nu a} + h_{\nu a}h_{\mu a} = 2g_{\mu\nu} \quad \text{by (1).}$$

---

[*] Flint, 'Proc. Roy. Soc.,' A, vol. 121, p. 676 (1928).

[†] 'Z. Physik,' vol. 50, p. 336 (1928).

The same theory applies to the relation

$$\gamma_\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2\delta_\mu{}^\nu .$$

This corresponds to the relation $h_{\mu a} h^{\nu a} = \delta_\mu{}^\nu$, and the operator form is obtained from (3) by multiplying by $g^{\lambda\nu}$.

The other operator of importance is one which is antisymmetrical in μ and ν, viz.,

$$s^{\mu\nu} = \tfrac{1}{2}\,(\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu).$$

In our notation this is

$$\tfrac{1}{2}\,(h^{\mu a} h^{\nu b} - h^{\nu a} h^{\mu b})\varepsilon_a \varepsilon_b . \tag{4}$$

If this operates scalarly on $h_{\lambda o}\varepsilon_o$ we obtain

$$s^{\mu\nu} \,.\, h_{\lambda o}\varepsilon_o = \tfrac{1}{2}\delta^\nu{}_\lambda h^{\mu a}\varepsilon_a - \tfrac{1}{2}\delta^\mu{}_\lambda h^{\nu a}\varepsilon_a .$$

There is thus a remarkable resemblance between the notation and that of Schroedinger.

The Nature of the Problem.

The actual character of the problem appears to be the discovery of the modification necessary to the Riemannian variation

$$\frac{\partial A^\mu}{\partial x^\nu} + \Gamma_{\sigma\nu}{}^\mu A^\sigma . \tag{5}$$

It appears that the successful inclusion of quantum mechanics into a unified scheme depends on the addition of a quantity to $\Gamma_{\sigma\nu}{}^\mu$ and, in our view, upon the extension to a five dimensional continuum.

One of the methods adopted* is to write

$$\frac{\partial A^\mu}{\partial x^\nu} + (\Gamma_{\sigma\nu}{}^\mu + S_{\sigma\nu}{}^\mu)\, A^\sigma \tag{6}$$

instead of (5); in which $S_{\sigma\nu}{}^\mu$ is related to Π^μ, where Π^μ is the generalized momentum.

In the paper referred to it is shown that this depends upon the adoption of a metric and the metric is determined by the equations of Dirac.

Both Fock and Schroedinger find a modification to the Riemannian metric. For Schroedinger's modification we find that the covariant matrices undergo the variation

$$\frac{\partial \gamma_\iota}{\partial x^l} = \Gamma_{\iota l}{}^\mu \gamma_\mu + \Gamma_l \gamma_\iota - \gamma_\iota \Gamma_l . \tag{7}$$

---

* Fisher and Flint, 'Proc. Roy Soc.,' A, vol. 126, p. 644 (1930).

This takes the place of

$$\frac{\partial \gamma_\iota}{\partial x^l} = \Gamma_{\iota l}{}^{\mu}\,\gamma_\mu, \tag{8}$$

which would lead to a contradiction if it were applied to all points in the neighbourhood of the point at which γ_ι is considered, for $\delta\gamma_\iota$, from (8), is a perfect differential only if the curvature vanishes at the point considered.

Schroedinger* obtains (7) from the assumption that the actual variation $\delta'\gamma_\iota$ differs from that given by (8) only by an infinitely small transformation.

In the Einstein system, which we adopted in the paper quoted, the law of variation is

$$\frac{\partial A^\mu}{\partial x^\nu} + \Delta_{\epsilon\nu}{}^{\mu} A^\epsilon \tag{9}$$

for a contravariant vector, and

$$\frac{\partial A_\mu}{\partial x^\nu} - \Delta_{\mu\nu}{}^{\epsilon} A_\epsilon \tag{10}$$

for a covariant vector.

The difficulty mentioned above does not arise, for in this notation, during a parallel displacement about a closed path, a vector returns to its original value and position. The quantity corresponding to curvature in this case vanishes and there is no difficulty in adopting (9) and (10) as our laws of variation.

The value of $\Delta_{\epsilon\nu}{}^{\mu}$ in terms of the orthogonal components is

$$h^{\mu a}\,\frac{\partial h_{\epsilon a}}{\partial x^\nu}\,.$$

The covariant derivative of $h_{\mu a}$ vanishes in this system, as may be readily seen by replacing A_μ by $h_{\mu a}$ in (10). But, in order to compare the laws (9) and (10) with (6) and (7), we must introduce the Christoffel symbol $\Gamma_{\epsilon\nu}{}^{\mu}$.

Now it can be shown that

$$\Gamma_{\epsilon\nu}{}^{\mu} = \Delta_{\epsilon\nu}{}^{\mu} - \Lambda_{\epsilon\nu}{}^{\mu} + \Theta_{\epsilon\nu}{}^{\mu}, \tag{11}$$

where

$$\Lambda_{\epsilon\nu}{}^{\mu} = \tfrac{1}{2}(\Delta_{\epsilon\nu}{}^{\mu} - \Delta_{\nu\epsilon}{}^{\mu})$$

and

$$\Theta_{\epsilon\nu}{}^{\mu} = g^{\mu\rho}\,(g_{\epsilon\sigma}\,\Lambda_{\rho\nu}{}^{\sigma} + g_{\nu\sigma}\,\Lambda_{\rho\epsilon}{}^{\sigma}).$$

Thus (10) may be written in the form

$$\frac{\partial A_\mu}{\partial x^\nu} - \Gamma_{\mu\nu}{}^{\epsilon}\,A_\epsilon - (\Lambda_{\mu\nu}{}^{\epsilon} - \Theta_{\mu\nu}{}^{\epsilon})\,\Lambda_\epsilon, \tag{12}$$

---

* *Loc. cit.*, p. 108, equation (7).

or writing

$$\mathrm{T}_{\mu\nu}{}^{\epsilon} = \Lambda_{\mu\nu}{}^{\epsilon} - \Theta_{\mu\nu}{}^{\epsilon},$$

$$\frac{\partial \mathrm{A}_\mu}{\partial x^\nu} - \Gamma_{\mu\nu}{}^{\epsilon}\mathrm{A}_\epsilon - \mathrm{T}_{\mu\nu}{}^{\epsilon}\mathrm{A}_\epsilon. \tag{13}$$

Thus we have added another term, $\mathrm{T}_{\mu\nu}{}^{\epsilon}\,\mathrm{A}_\epsilon$, to the Riemannian variation. But from the vanishing of the covariant derivative of $h_{\mu a}$ we have

$$\frac{\partial h_{\mu a}}{\partial x^\nu} = \Gamma_{\mu\nu}{}^{\epsilon} h_{\epsilon a} + \mathrm{T}_{\mu a}{}^{\epsilon} h_{\epsilon a}. \tag{14}$$

This is our form of Schroedinger's variation of γ_ϵ (7).

The only difference between (13) and (6) is the replacement of S by T and the procedure to the quantum equations is similar in the two cases.

The difference is in the underlying metrical interpretation and the choice of one or the other must depend upon developments.

In both cases we again have evidence that quantum phenomena are the physical expression of geometry or of metric

The resemblance between our present notation and that of Schroedinger, and the fact that McVittie has brought forward some evidence in favour of Schroedinger's generalization, seems to indicate that possibly the notation used here is to be preferred to that we used previously

It is interesting to note how the unearthing of this geometrical significance comes about.

In former papers we have seen that it is a consequence of the discovery of the first order equations from which the wave equation may be derived.

In making use of orthogonal components it is a consequence of finding the first order expression of the line element from which $ds^2 = g_{mn}\,dx^m\,dx^n$ is derived.

We regard ds as having orthogonal components, of which ds_a is typical Then $ds^2 = \Sigma\, ds_a{}^2$, and in terms of the h's, $ds_a = h_{\mu a}\,dx^\mu$. Hence

$$ds^2 = h_{\mu a}\,h_{\nu a}\,dx^\mu\,dx^\nu = g_{\mu\nu}\,dx^\mu\,dx^\nu.\text{*}$$

Thus the linear form of ds introduces the h-components and, in imposing a limitation upon these components, we are able to derive the wave equation in a very simple way.†

The quantum theory indicates that a special case of Einstein's theory of parallelism is adopted in nature. The case is somewhat similar to the limitation in the value of the tensor $G_{\mu\nu}$ in the theory of relativity.

---

* *Cf.* Fock and Iwanenko, 'Z. Physik,' vol. 54, p. 798 (1929).

† Flint, 'Proc. Roy. Soc.,' A, vol. 121, p. 676 (1928).

The First Order Equations of the Quantum Theory.

It has been shown, Fisher and Flint *loc. cit*, that Dirac's equations can be derived from a set of equations expressing the vanishing of a divergence. The similarity with the present case leads to a simple form of these equations in the present notation.

We conclude that $T_{\mu\nu}{}^{\epsilon}$ is related in the same way as $S_{\mu\nu}{}^{\epsilon}$ to the generalized momentum and that here, as in the former case, the quantity Π replaces ϕ, when the five dimensional system is employed.

Thus extending the previous work (Flint, *loc. cit.* (1928)) we write

$$\Lambda_{s\nu}{}^{\nu} \propto \Pi_{s}. \tag{15}$$

With this limitation the quantum equations become the expression of the vanishing of the divergence of a tensor density.

A tensor density is defined here as a tensor multiplied by h', and is similar to the same quantity in the theory of relativity, h' being the value of the determinant $|h_{\mu a}|$, composed of the $h_{\mu a}$'s. The vanishing of the divergence of a vector, A^{μ}, gives the equation

$$\frac{\partial A^{\mu}}{\partial x^{\mu}} - \Delta_{\epsilon\mu}{}^{\mu} A^{\epsilon} = 0, \quad \text{by (9).}$$

Writing this in the form in which $(h'A^{\mu})$ occurs as the varying quantity, we obtain

$$\frac{\partial}{\partial x^{\mu}} (h'A^{\mu}) + \Delta_{\epsilon\mu}{}^{\mu} (h'A^{\epsilon}) - A^{\mu} \frac{\partial h'}{\partial x^{\mu}} = 0.$$

But

$$\frac{\partial h'}{\partial x^{\mu}} = h' \Delta_{\sigma\mu}{}^{\sigma}.$$

Hence we obtain

$$\frac{\partial}{\partial x^{\mu}} (h'A^{\mu}) - (\Delta_{\sigma\mu}{}^{\sigma} - \Delta_{\mu\sigma}{}^{\sigma}) h'A^{\mu} = 0, \tag{16}$$

or

$$\frac{\partial}{\partial x^{\mu}} (h'A^{\mu}) - 2\Lambda_{\sigma\mu}{}^{\sigma} (h'A^{\mu}) = 0, \tag{17}$$

or

$$\frac{\partial}{\partial x^{\mu}} (h'A^{\mu}) - \frac{2\pi i}{h} \Pi_{\mu} (h'A^{\mu}) = 0, \tag{18}$$

where we now write $\Lambda_{\sigma\mu}{}^{\sigma} = \frac{\pi i}{h}\Pi_{\mu}$, as an extension of the earlier value, and in agreement with the views expressed in the paper quoted (Fisher and Flint, *loc. cit.*, equations (15) and (22)).

Equation (17) introduces us to the operator $\left(\frac{\partial}{\partial x^{\mu}} - 2\Lambda_{\sigma\mu}{}^{\sigma}\right)$, which is similar in form to that introduced by Schroedinger and appears to be of special importance in the theory. It reminds us of the operator, $u_n = \frac{h}{2\pi i}\left(\frac{\partial}{\partial x^n} - e\phi_n\right.$ which plays an important part in the wave equation, and of which it is a generalization. The equation (18) is somewhat similar to the set of equations obtained in the earlier paper, to which we have just referred. But the difference between the vector A^{μ} and the tensor $A^{\mu\nu}$ must be noted.

This seems to indicate that the quantity (h' A^{μ}) should be replaced by $h'A^{\mu\nu}$ and that we require the divergence of this tensor density.

The vanishing of this divergence gives

$$A^{\mu\nu}\,;\,\mu - 2\Lambda_{\alpha\sigma}{}^{\sigma}A^{\alpha\nu} = 0, \tag{19}$$

where the first term denotes the covariant derivative of $A^{\mu\nu}$. If we write $A^{\mu\nu}$ in the form $h^{\mu a}h^{\nu b}\psi_{ab}$ we obtain from (17)

$$h^{\nu b}h^{\mu a}\left(\frac{\partial}{\partial x^{\mu}} - 2\Lambda_{\mu\sigma}{}^{\sigma}\right)\psi_{ab} = 0 \tag{20}$$

This equation is now similar to that previously obtained and is remarkably like that of Schroedinger. From this point the argument proceeds as before.

The Union of Cosmic and Quantum Phenomena.

Eddington* has deduced a relation between the radius of curvature, R, of the universe, and quantities associated with atomic phenomena. The relation is

$$\frac{n^{\frac{1}{2}}}{R} = \frac{mc^2}{e^2},$$

n denoting the number of electrons in the universe.

This result has been shown to fit in with an identity previously adopted by us.† The radius of curvature appears very naturally in Schroedinger's work and is related to the additional term in his law of variation of the matrices (7) (*cf.* also his paper, equation (9)).

---

* 'Proc. Roy. Soc.,' A, vol. 133, p. 605 (1931).

† Fahmy, 'Proc. Phys. Soc.,' vol. 45, p. 67 (1933).

In exactly the same way our equation (13) gives a relation between the Riemannian curvature and the quantities $T_{\mu\nu}{}^{s}$, which, through $\Lambda_{\mu\nu}{}^{s}$, are related to the generalized momentum.

Summary.

The union of Dirac's equations with the theory of relativity is discussed by the use of Einstein's theory of parallelism.

The relation of the present work to that of Schroedinger is shown and some difficulties in Schroedinger's theory are avoided.

A feature of this paper is the natural introduction of a relation between the Riemannian curvature and quantities characteristic of the quantum theory—a feature which the paper has in common with the work of Schroedinger.

Experiments on the Protons produced in the Artificial Disintegration of the Nitrogen Nucleus.

By E. C. Pollard, Ph.D., Assistant Lecturer in Physics, University of Leeds.

(Communicated by R Whiddington, F R.S.—Received February 18, 1933.)

1. *Introduction.*

It is now familiar that the events relative to the disintegration of radioactive bodies, to the scattering of alpha-particles and to artificial disintegration can be explained most simply by supposing the nucleus to have a potential field similar to a volcano and crater, whose walls can be penetrated by the alpha-waves. Using this idea to describe artificial disintegration where an alpha-particle hitting a nucleus causes a proton to be thrown out, we are led to the question of the fate of the alpha-particle—does it enter the crater and stay inside or does it re-emerge or does it cause the ejection of the proton from outside ? For the element nitrogen we have the direct experimental answer from Blackett's work, where the tracks of the disintegration process are seen in an expansion chamber . all the observations show only three tracks, a fact which suggests that the alpha-particle stays inside the disintegrated nucleus ; other less direct evidence renders it likely that this capture of the alpha-particle always takes place.

The alpha-particle can make its entry either by surmounting the top of the potential barrier and falling inside; or by passing through the barrier. The calculation of the probability of this latter occurrence shows that except very near the top of the barrier it is extremely rare unless the energy of the incident particle coincides with an unoccupied level of energy inside, when the probability of entry rises enormously. This resonance was first shown to occur for the element aluminium by Pose.* The object of the following experiments was to see whether the alpha-particle entered by resonance or over the top of the barrier, and if in the former way to fix the energy of the internal resonance level; if in the latter, to find the height of the potential barrier. In addition it was intended to investigate whether the protons produced from nitrogen are strictly homogeneous (that is, whether they have the same energy, a spread of energies, or two discrete energy values) by finding the absorption curve at smaller values than had previously been attempted Such experiments might give confirmation to the suggestion put forward by Blackett and Lees,† that there is a selective absorption of the alpha-particles at about 2·6 cm. range.

2. *Experiments on the Method of Entry of the Alpha-particle*

There is an important difference between the conditions of entry by resonance and by crossing the top of the barrier Entry by resonance takes place for a narrow band of energies *above and below which* the alpha-particle cannot make its way in: Chadwick and Constable‡ have determined the width of these bands for the element aluminium, they are roughly 250,000 electron volts. Entry over the top is possible so long as the alpha-particle has an energy greater than the energy of the top of the potential barrier. Thus limitation to the entry is *two sided* for resonance; *one sided* for scaling the barrer. This will be made clear from fig. 1 where the probability of entering is shown on the right and half the potential barrier drawn on the left.

If now we bombard a narrow layer of nitrogen gas with a source of alpha-particles of constant energy, and arrange to vary the density (and so the absorption) of the layer, we shall vary the limits of energy of the disintegrating alpha-particle. Thus with a layer of nitrogen at atmospheric pressure 4·1 cm. thick and a source of polonium (range 4·1 cm. in nitrogen) there will be alpha-particles of all energies from zero to the maximum encountering nitrogen molecules. On reducing the pressure the smaller energies do not encounter nitrogen molecules since the alpha-particles have traversed the layer before their velocity is reduced to the lower value, and finally when the pressure is very low only the maximum energy alpha-particles have any effect. Then if there is resonance we should expect such alpha-particles to have too great an energy and so to exert no disintegrating action, while for entry over the top we should expect disintegration to occur. Now if we increase the pressure we shall expect the yield to rise suddenly at a definite value if entry is by resonance (where the alpha-particle energy is reduced to the critical value) and to reach a maximum quickly, while for entry over the top there will be a steady rise from the lowest pressures to a flat maximum. The two types of curve are indicated in fig. 2.

---

* 'Z. Physik,' vol. 64, p. 1 (1930).

† 'Proc. Roy. Soc.,' A, vol. 136, p. 337, and plate 7.

‡ 'Proc. Roy. Soc.,' A, vol. 135, p. 48 (1932).

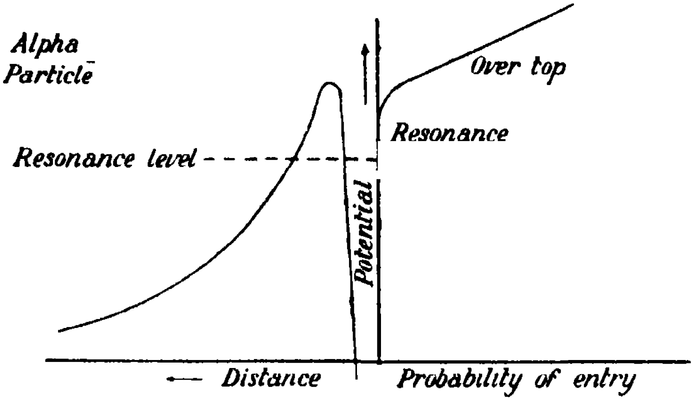


FIG. 1.

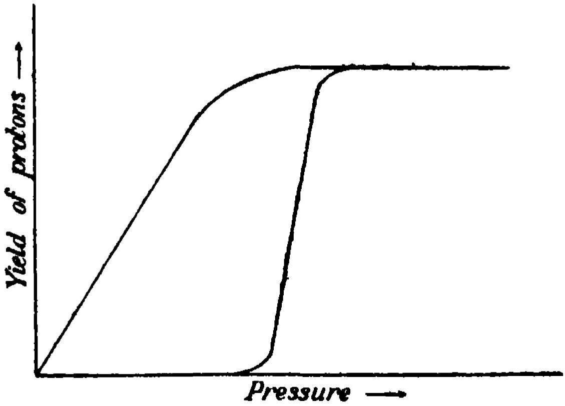


FIG. 2.

Further evidence can be obtained from the shape of the absorption curve for the protons produced from a thick layer of gas. An absorption curve tells us the spread of energies of the protons; a proton of definite energy has a definite range and any straggling at the end of a group of protons means that the energies are distributed over a region of values. There is reason to believe that the energy change in transition from one nucleus to the other is constant and so any variation in proton energy must mean a variation in the energy of the alpha-particle producing it, *i.e.*, for " resonance " the proton will have one energy only, while for entry over the top the protons will have values spread between two limits (one due to the maximum energy of the alpha-particle, the other to the energy value of the top of the barrier). The absorption curve will thus have a definite " tail." From the length of this tail we can calculate the height of the barrier.

These considerations indicated that the previous experimental work admitted of explanation in different ways. Rutherford and Chadwick,* investigating the variation of yield of protons with increasing alpha-particle energy found a steady increase beyond about 2·8 cm. range, which points to entry over the top and indicates the height of the potential barrier. Comparison with the evidence available for other light elements from scattering experiments tends to confirm this. On the other hand Chadwick, Constable and Pollard,† working with a thin layer of nitrogen, suggested a definite group of protons. This could be taken to be a resonance group. The two experiments do not cover the same ground and it was felt that additional work was worth while in order to link the two. This work has shown that the two experiments are quite consistent, and that they contain no contradiction.

3. *Experiments with a Layer of Variable Density.*

The arrangement of the apparatus can be seen from fig. 3. Alpha-particles from the source A are absorbed either in the gas or in the gold foil closing the glass tube at B.

The protons produced in the disintegration traverse the gold foil and cause ionization in the chamber which is amplified by valves and detected either by telephone or by a loudspeaker unit which rotates a small mirror and makes a spot of light record on a photographic strip. Various absorptions can be placed in the path of the protons by means of mica screens covering the entrance to the ionization chamber. The experiment consists of varying the pressure of the gas (air was used since oxygen does not disintegrate) and plotting the yield of protons against the pressure; as explained above this should decide the method of alpha-particle entry. Details of the preparation of the source have been given in a previous paper* in which will also be found a preliminary account of the varying pressure experiment. In such work it is important to ensure that the alpha-particles themselves have a uniform velocity, otherwise the sharpness of the yield at resonance will be diminished. Actually the source was prepared by electrolytic deposition and a test showed that two-thirds of the alpha-particles had ranges between 3·7 and 3·9 cm.

---

* " Radiations from radioactive substances," p. 297.

† ' Proc. Roy. Soc.,' A, vol. 130, p. 463 (1931).

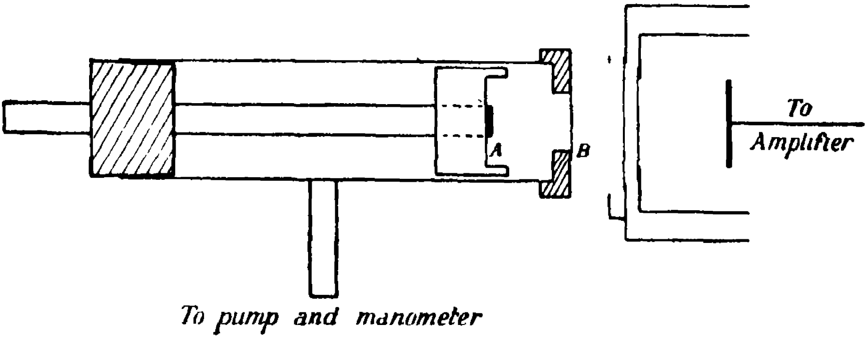


FIG. 3.

The distance AB was 11 mm., the opening B was 14 mm. in diameter covered by gold foil of stopping power 3·6 cm., opening of counter 10 mm. radius, distance from it to B, 16 mm. The airtightness of the space was ensured by a thin piece of mica of 2 cm. air equivalent waxed outside the gold.

The results of counting the protons are shown in fig. 4, each curve corresponding to a different absorption in the path of the protons; stopping powers were found from mica screens and gold foil. The mica screens from the weight per square centimetre (1·43 mg. for 1 cm. air equivalent), and the gold by direct measurement using an alpha-particle source. Random fluctuations account for the fact that the points do not all lie smoothly on the curve, particularly in the 9·3 cm. curve where only a hundred particles were counted at each point. The abscissæ are given in centimetres of mercury and the corresponding minimum range is put below each value. It will be seen at once that the curves are of the type expected if *entry is over the top* and *not by resonance*. This is therefore the first conclusion to be drawn from the experiments.

---

* Pollard, 'Proc. Leeds Phil. Soc.,' vol. 2, p. 324 (1932).

4. *Absorption Curve Experiments.*

For the purpose of plotting an absorption curve the distance AB was increased to 18 mm. and a thinner gold foil, stopping power 2·4 cm., and thinner mica used to close the end: this necessitated a smaller opening, and a cover with perforations was substituted for the single hole previously used. Allowing for the absorption in the gas itself the smallest range that could be measured was 5·6 cm. The plotting of this curve is complicated by the presence of the natural H-particles which are always given off by any source: a separate curve for these was plotted with the space evacuated and their numbers deducted from those of the actual count with gas present. Very great care was also taken to eliminate any hydrogen impurity in the air and elaborate drying arrangements were used—as a routine the air was dried by standing over strong sulphuric acid for at least 2 days and then passed through sulphuric acid and two stages of phosphorus pentoxide on its way to the disintegration space. Control experiments with methane showed that 6 cm. water vapour pressure would be needed to account for the yield obtained, an amount which is certainly far in excess of that present in the actual air used. At these small absorptions telephone counting was abandoned, reliance being placed only on the photographic records, which were separately counted by two persons.

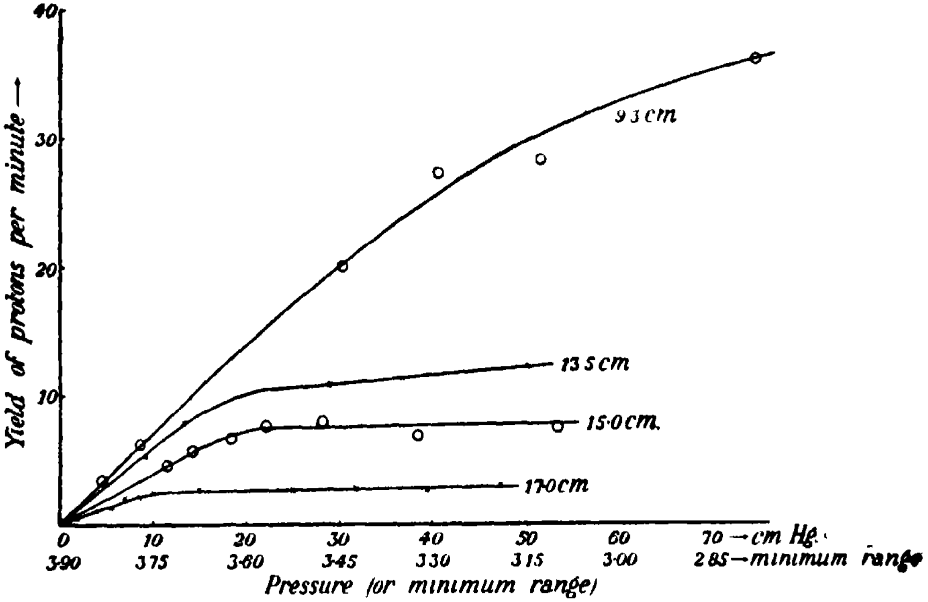


Fig. 4.

The absorption curve obtained is shown in fig. 5. For the purposes of ready reference the curves found by Chadwick, Constable and Pollard (*loc. cit.*) and Steudel* are given on the same graph—it will be seen that the present work agrees well with that of Steudel, as far as it is taken, but apparently contrasts sharply with the other curve. This can be simply explained, since the work in Cambridge was made with a relatively thin layer of nitrogen (3 mm.) for which the available energies of alpha-particle would be considerably reduced in extent, so producing an apparently homogeneous group—in fact, one whose properties would fit with resonance entry through a band of energies from 3·6 to 3·9 cm. alpha-particle range. There is no contradiction between the two experiments—the value of the work by Chadwick, Constable and Pollard is that it shows the nuclear energy change to be constant, knowledge which is necessary for the explanation of both the other experimental results.

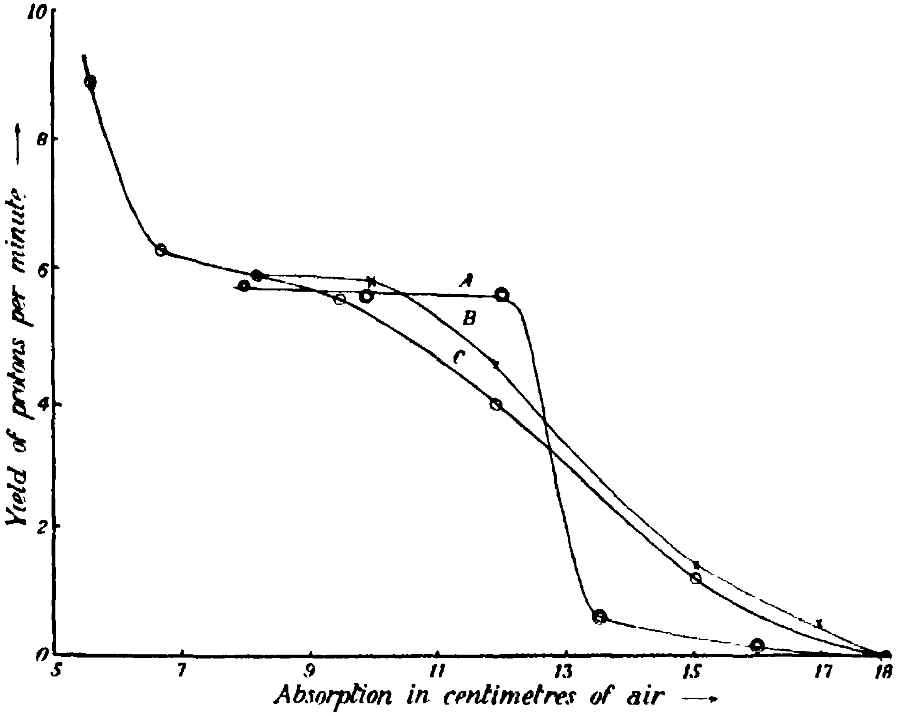


Fig. 5.—Curve A, Chadwick, Constable and Pollard ; B, Steudel ; C, Pollard.

The difference between the present curve and that of Steudel can most probably be put down to difference of geometrical conditions : the theoretical curve for protons in one direction only can never be attained; in practice a certain solid angle must be chosen and the greater this is, the less definitely does any critical position on the curve show itself. It is difficult to estimate the effect of the geometrical conditions with certainty. The protons can be taken as coming from the middle of the gas, which means they subtend an outside angle of 22° at the counter. This introduces an average error of about 8% in the ranges observed.

---

* 'Z. Physik,' vol. 77, p. 148 (1932).

If we assume that the flattening of the curve between 9 and 10 cm. is due to the inability of the alpha-particles to surmount the potential barrier then we can calculate the height of the barrier We assume that the atomic energy change Q is a constant.

The energy and momentum equations give*

$$2m_nQ = m_pv_p^2(m_p + m_n) - MV^2(m_n - M) - 2MVm_pv_p,$$

where m_n is the mass of the new nucleus.

m_p is the mass of the proton.

M is the mass of the alpha-particle.

v_p is the velocity of the proton.

V is the velocity of the alpha-particle.

For the maximum range we know all except Q which is therefore determinable. We can then apply this value of Q to find the value of the velocity of the alpha-particle which gives a proton of velocity corresponding to 9 5 cm. range. The energy of this alpha-particle gives the height of the potential barrier. We find the value $4\cdot1 \times 10^6$ *electron volts.*

If we follow the curve closer in there is a rise at 6·6 cm. range. At these small absorptions there is difficulty in counting for two reasons: first, because we are concerned with fast protons which produce relatively small ionization and so are not easy to detect; and second, because the natural H-particles produced by the source are greater in numbers at these short ranges, so that there is a little uncertainty about very short range measurements. These natural H-particles show a sloping straight line absorption curve, roughly a fifth of the nitrogen yield. Nevertheless, the rise in yield, which is confirmed by a rise in the number of large size deflections on the records, indicates that there is present a *second group* of protons of range 6·6 cm.

The possibility of the existence of a second group has been suggested by Blackett and Lees (*loc. cit.*) who found, in expansion chamber experiments, that alpha-particles of ranges 24, 23 and 29 mm. could produce disintegration.

---

* Chadwick, Constable and Pollard (*loc. cit.*).

This corresponds to a group of protons of shorter range, due to entry by resonance. Since nitrogen does not emit an artificial gamma-ray like fluorine or aluminium it is not possible to explain this second group in terms of two nuclear levels, one excited and one normal, to which the nucleus reverts from the excited state; if the alpha-particle is captured the only explanation is the existence of a virtual alpha-particle level below the top of the barrier which will account for the presence of shorter range protons. The truth of this resonance can be confirmed by repeating the varying pressure experiment at small absorptions to see whether there is a rise in yield when the minimum energy of the alpha-particle falls to the value of the virtual energy level. In addition, the flattening of the curve plotted for these small absorptions will tell us the height of the potential barrier, deduced from the least energy of the alpha-particle which will be effective. The curve resulting from this experiment is shown in fig. 6. The circles are points taken with 5·6 cm. absorption,

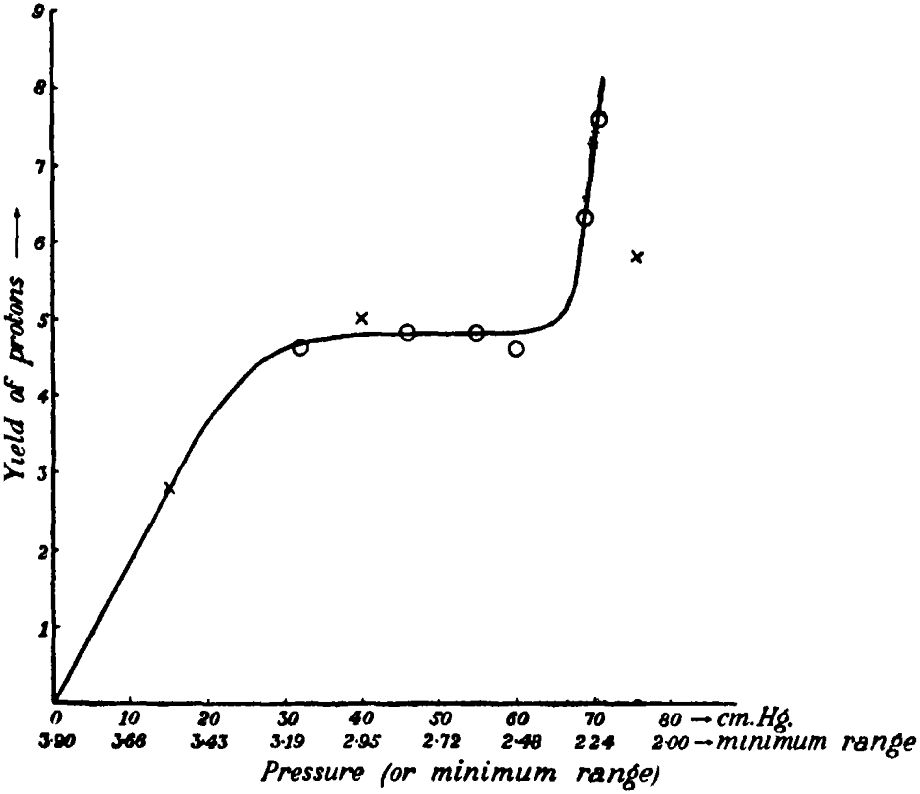


Fig. 6.

the crosses at 7·2 cm. It will be seen that there is a flattening when the minimum range of the alpha-particle is 3·00 cm.; alpha-particles of lower energy than this do not cause disintegration, presumably because their energy is insufficient to cross the barrier, whose height is therefore the energy of a 3·00 cm. alpha-particle or 4·36 × 10^6 *electron volts*. This agrees satisfactorily with the value 4·1 × 10^6 previously found.

There is a definite rise in yield when the minimum range of the impinging alpha-particle falls as low as 2·2 cm. and this is a confirmation of the second group found in the absorption curve experiments. This rise in yield is not found for the protons of 7·2 cm. range. This experiment therefore fixed the virtual alpha-particle level at 2·2 cm. and the absorption curve gives the range of the protons produced as 6·6 cm.; we can therefore use the energy and momentum equations to calculate the value of Q, the nuclear energy change, which should agree with the value found from the maximum range. Substituting the values in the equation already quoted we find − 1·32 × 10^6 electron volts, while the values for maximum range give − 1·26 × 10^6 electron volts, an agreement which is well within the errors of experiment.

In these calculations a knowledge of the relation between the range and the velocity of the proton is necessary. The values given by Blackett* have been plotted on a curve from which the particular velocities have been read off.

Discussion.

In these experiments an attack has been made on the problem of the minimum energy of an alpha-particle which will cause disintegration. In view of the potential barrier conception of a nucleus, such a problem has no meaning unless it is understood whether the alpha-particle, which enters a nucleus, does so over the top of the barrier or by resonance through the barrier, and the first aim was therefore to settle this point. It is shown that entry is over the top and the height of the barrier is fixed.

This height agrees with the results of the earlier experiment of Rutherford and Chadwick, who found that the yield of protons ceased when the energy of the incident alpha-particle fell below 2·8 cm. range, which gives a slightly lower figure for the height of the nuclear barrier than is found in the present work.

Now that we know the entry is in this manner we can adopt the simple explanation of the tail at the end of groups of protons from elements where the barrier is low, such as boron or nitrogen; this tail is due to the production

* 'Proc. Roy. Soc.,' A, vol. 135, p. 132 (1932).

of protons by alpha-particles with energies ranging from that of the top of the potential barrier to the maximum energy of the radio-active particle itself. Simple calculation of this tailing off gives the value of the height of the barrier.

Stress has been laid on the fact that most elements that produce protons under alpha-particle bombardment also produce artificial gamma radiation and this has been related to the presence of groups of protons.* Nitrogen does not give any gamma radiation and this can be understood if there is only one nuclear level possible for the alpha-particle; this may be due to the fact that the alpha-particle level is higher than the proton level. These experiments, which indicate a second group of protons, are therefore most readily explained in terms of resonance (while other elements such as aluminium or fluorine offer the possibility of different nuclear levels). This is the explanation suggested by the varying pressure experiment which gives directly the value of the resonance level, and enables an independent determination of the nuclear energy change to be made. There is, of course, the possibility of other resonance levels existing at lower energies; to find these it will be necessary to observe the protons in the backwards direction since the range of the emitted protons will be less than that of the alpha-particles from the source.

It was suggested by Urey† as a result of an analysis of the cloud chamber tracks photographed by Blackett and Harkins, that the oxygen nucleus may not always have the same mass, *i.e.*, that groups should exist due to different nuclear energies. These experiments give no confirmation to this suggestion.

In conclusion I should like to thank Professor Whiddington for his interest and encouragement, and for giving me facilities for this research. I am grateful to Dr. Chadwick and also Professor Sidney Russ of the Middlesex Hospital, for their gift of radon tubes, and to Dr. Shirodkar for assistance with the counting.

Summary.

Experiments to determine the manner of entry of the alpha-particle into the nitrogen nucleus are made; these indicate that entry is in general over the top of the barrier and the height of this barrier is fixed as between 4·1 and 4·4 × 10^6 electron volts.

Further investigation of the absorption curve of the protons confirms the work of Steudel on this element and gives an indication of a second group at 6·6 cm. range, due to resonance with an alpha-particle of 2·2 cm. range or 3·5 × 10^6 electron volts energy.

---

* Chadwick, Constable and Pollard (*loc. cit.*).

† 'Phys. Rev.,' vol. 37, p. 923 (1932).

Ionization of Mercury Vapour by Positive Ions of Mercury and Potassium.

By Rafi Mohammed Chaudhri, M.Sc. (Alig.), Ph.D. (Cantab.).

(Communicated by Lord Rutherford, O.M., F.R.S.—Received February 24, 1933.)

It was postulated by Townsend some thirty years ago that ionization of gases by positive ions played an important part in a self-maintained electric discharge through a gas. There was, however, no direct experimental evidence in favour of this hypothesis.

Very few experiments have been done so far in connection with the solution of this problem and it appears that no satisfactory method has been devised to investigate it. The experiments of Sutton* which showed that positive ions of potassium of energy above 100 volts ionized neon and argon are not conclusive, since the results obtained can be explained in terms of the secondary effects caused by the impact of primary positive ions and excited atoms upon different parts of the apparatus.

When positive ions of a few hundred volts energy bombard metal surfaces, they are able to liberate secondary electrons some of which possess energy as high as 20 volts, as has been shown by Oliphant.† This energy is greater than the ionization energy of several gases, and no allowance has been made for this fact in these experiments. Similar uncertainties exist in the experiments of Sutton and Mouzon,‡ Beeck and Mouzon,§ Mouzon,|| and Brasefield,¶ and it is by no means safe to conclude, on the basis of these experiments, that positive ions ionize gases directly.

The author has investigated the ionization of mercury vapour by Hg^+ and K^+ ions, and the method employed is such that the secondary effects of the positive ions and the ionization produced by them can be separately distinguished.

The Theory of the Experiment.

Consider two parallel plates P_1 and P_2, fig. 1, at some distance apart. The upper plate has a rectangular narrow slit S_2 in the centre, and the space between

---

* 'Phys. Rev.,' vol. 33, p. 364 (1929).

† 'Proc. Roy. Soc.,' A, vol. 127, p. 373 (1930).

‡ 'Phys. Rev.,' vol. 35, p. 694 (1930); vol. 37, p. 319 (1931).

§ 'Phys. Rev.,' vol. 38, p. 967 (1931).

|| 'Phys. Rev.,' vol. 41, p. 605 (1932).

¶ 'Phys. Rev.,' vol. 42, p. 11 (1932).

the plates is filled with the gas to be ionized. If a beam of positive ions is sent through the centre of the plates, the direction of the beam being perpendicular to the plane of the paper and parallel to the length of the slit, and if positive ions cause ionization of the gas, electrons will be produced in the path of the beam at O. If a potential V is applied between the two plates making P_1 positive with respect to P_2 then the electrons produced in the path of the beam will be pulled upward, and will shoot out of the rectangular slit with energy V/2 volts. On the other hand, the secondary electrons liberated from the bottom plate due to the bombardment by scattered positive ions or other agencies like metastable or excited atoms* will come out of the slit S_2 with energy of very nearly V volts. In the presence of the electrostatic field alone, the two groups of electrons will emerge normally to the plate P_1.

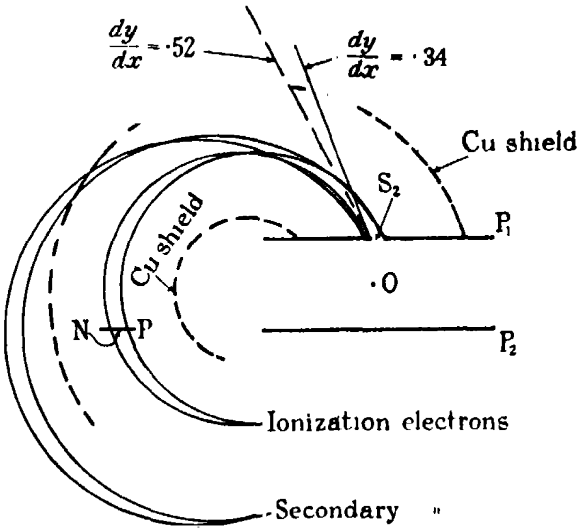


FIG. 1.

If now a magnetic field H be simultaneously applied, parallel to the beam of positive ions, *i.e.*, perpendicular to the plane of the paper, the electrons of energies V and V/2 volts will describe circular paths of different radii depending upon the field strength H. The electrons produced through ionization can therefore be separated from the secondary electrons.

It should be pointed out here that under the simultaneous action of the electrostatic and magnetic fields the electrons will describe a part of a cycloid between the plates and will therefore emerge inclined at certain angles to the plate and not perpendicular to it.

---

* Oliphant, 'Proc. Roy. Soc.,' A, vol. 124, p. 228 (1929).

The position of an electron of mass " m " and charge " e " under the simultaneous action of an electrostatic field X and a magnetic field H acting perpendicular to each other between the plates along the axes of x and z respectively is given at any time " t " by

$$y = \frac{X}{\omega H}(\omega t - \sin \omega t), \tag{1}$$

$$x = \frac{X}{\omega H}(1 - \cos \omega t). \tag{2}$$

This is the equation of a cycloid; ω is equal to $H \, . \, e/m$ and the axis of y is perpendicular to the axes of x and z.

From equations (1) and (2)

$$\frac{dy}{dt} = \frac{X}{\omega H}(\omega - \omega \cos \omega t), \tag{3}$$

and

$$\frac{dx}{dt} = \frac{X}{\omega H}(\omega \sin \omega t). \tag{4}$$

Therefore

$$\frac{dy}{dx} = \frac{1 - \cos \omega t}{\sin \omega t}. \tag{5}$$

From equation (2), for $x = a$, if $t = t'$, then

$$a = \frac{X}{\omega H} 2 \sin^2 \frac{\omega t'}{2},$$

$$\sin \frac{\omega t'}{2} = \sqrt{\frac{\omega H a}{2X}},$$

$$= \sqrt{\frac{H}{2} \cdot \frac{e}{m} \cdot \frac{H}{X} \cdot a}. \tag{6}$$

Consider a particular case in which the plates are at a distance of 2 cm. apart and a potential difference of 400 volts is applied between them. $X = 200$ volts per centimetre.

The ionization electrons start from the centre of the plates and therefore for them " a " is nearly 1 cm. From equation (6), for these electrons

$$\sin \frac{\omega t'}{2} = \frac{H}{200} \times 4{\cdot}1.$$

If $H = 16$ Gauss

$$\sin \frac{\omega t'}{2} = 0{\cdot}328, \textit{ i.e.}, \frac{\omega t'}{2} = 19° \, 12',$$

and

$$\frac{dy}{dx} = 0{\cdot}34. \tag{7}$$

Similarly it can be shown that for electrons starting from the bottom plate, with X equal to 200 volts per centimetre and " *a* " equal to 2 cm.

$$\frac{dy}{dx} = 0{\cdot}52. \tag{8}$$

We conclude therefore that the two sets of electrons after emerging through the slit S_2, will be inclined to the plate at angles whose tangents are 0·34 and 0·52. After emerging out of the space between the plates the electrons are under the action of the magnetic field alone and will describe circles of different radii.

An electron of energy V electron-volts, describes in a magnetic field H a circle of radius " *r* " given by

$$r = \sqrt{\frac{m}{e}} \times \frac{\sqrt{2V}}{H}. \tag{9}$$

For the electrons produced in the centre of the plates, in this particular case, V = 200 volts, and for the secondary electrons coming from the bottom plate, V = 400 volts. The values of " *r* " for the two groups, with H = 16 Gauss, are 3·02 cm. and 4·28 cm. respectively.

The paths of the two sets of electrons are shown for these particular conditions in fig. 1, the magnetic field and the primary ion current beam being perpendicular to the plane of the paper. It can be seen that the ionization and secondary electrons focus at two widely separated points. It may, however, be pointed out that the two sets of electrons can be made to describe approximately the same paths for very different potentials between the plates if $\sqrt{V/H}$ is kept constant, neglecting the small variation owing to their different inclinations to the upper plate with different X and H. It may also be observed that since the energies of the two groups of electrons are in the ratio of 1 : 2 the magnetic fields required to make them describe paths of the same radius will be in the ratio of $1 : \sqrt{2}$.

Apparatus.

A homogeneous beam of positive ions of mercury was produced by the method devised by Oliphant (*loc. cit.*, 1930). A hot cathode mercury arc at

100–120 volts, passing about 50 m.A. of current was established in the chamber "O" fig. 2.

The positive ions were pulled out of the ionized vapour by means of the negatively charged Langmuir probe K, of molybdenum, which had a small perforation in its centre. The positive ion beam which passed through the canal D was defined by a circular hole S_1, 1·5 mm. in diameter in the closed end of the copper cylinder F.

The parallel plates P_1 and P_2 were made of copper, 5·0 cm. × 5·0 cm. and were 2 cm. apart. They were held in position by tungsten seals spot welded to copper leads which were silver soldered to the plates. The upper plate had a rectangular slit S_2 in its centre, 3 cm. long and 5 mm. wide.

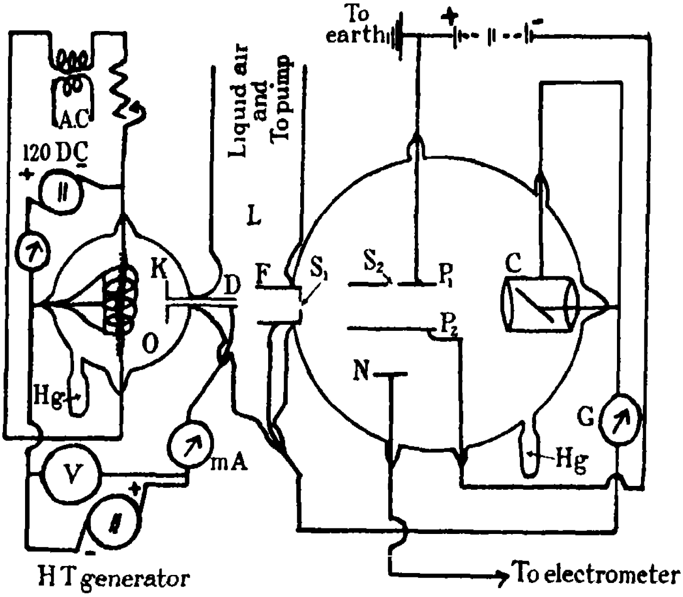


FIG. 2.

The centre of the two plates, the defining hole S_1 and the axis of the canal D were carefully adjusted to be in one straight line.

The primary beam of positive ions was collected in the Faraday cylinder system C.

A copper collecting plate N, 3 cm. long, was placed parallel to the slit S_2 at the point P, fig. 1, which was about 3 cm. away from the nearest edge of the bottom plate P_2. It was enclosed in a hollow copper channel which was screwed to the upper plate P_1 to shield the whole path of the electrons between the slit S_2 and the collector N, as shown by the dotted lines in fig. 1.

A uniform magnetic field over the path of the electrons was produced by a pair of large Helmholtz coils. Since the magnetic field required in the experiment was very small, all the metal parts were made of non-magnetic material to keep the field uniform at every point. The primary Hg^+ ion beam reaching the cylinder C was measured by a galvanometer of sensitivity corresponding to a deflection of 1 mm. in the scale for 10^{-10} amperes.

For a setting of the magnetic and electric fields which would bring electrons produced in the path of the beam which lies immediately beneath S_2 on to the collector N, the secondary electrons produced at the lower plate will collide with the copper shield. Secondary electrons which are not accelerated will be bent into circles of small radius by the magnetic field, while the efficient shielding ensures that no effects can be produced at N by diffusion of metastable atoms or any other agency. The secondary electrons emitted from the defining hole S_1, are not effective in producing ionization beneath the slit S_2, as they will be lost in the electrostatic field at the edge of the plates. The same argument holds for other stray electrons entering the space between the plates. The electron current to the collector N, measured by means of a quadrant electrometer, will therefore be a measure of the ionization of mercury vapour produced by positive ions between the plates.

The ionization chamber was connected to a reservoir of clean mercury, which provided the Hg vapour to be ionized. The experiment was done with Hg vapour at room temperature only.

The whole apparatus was made of pyrex. A liquid air trap was incorporated in it at L, fig. 2, to reduce the pressure of Hg vapour in the space between D and F. The apparatus was baked with a Bunsen flame and the experiment was performed in a vacuum produced by a two-stage " Gaede " steel diffusion pump.

In carrying out an experiment with K^+ ions, the Hg arc was replaced by a tungsten filament 1·0 cm. long and 2 mm. wide mounted on nickel leads exactly opposite the hole in K, and at a distance of 1·5 mm. from it. The filament was enclosed in a tubular glass chamber, to the side of which was connected a small tube containing distilled potassium, similar to the arrangement described by Moon and Oliphant.* By applying a negative potential to the probe K, which in this experiment had a smaller perforation of 1 mm. diameter, a homogeneous beam of K^+ ions of any desired energy could be obtained. The ion current at any particular energy was controlled by changing the vapour pressure of the potassium.

---

* 'Proc. Roy. Soc.,' A, vol. 137, p. 463 (1932).

A potential of about 80 volts was applied between the plates P_1 and P_2, and the strength of the magnetic field was varied from zero upwards. The copper shield was earthed and the whole ionization chamber was well shielded by thin tin foil on the outside of the glass. The electron current to the collector N was measured as the magnetic field H was varied.

If the direction of the magnetic field was reversed keeping other conditions unaltered there was no deflection of the electrometer other than a constant leakage current, so that the current reaching the collector in the previous experiment must have been an electronic current.

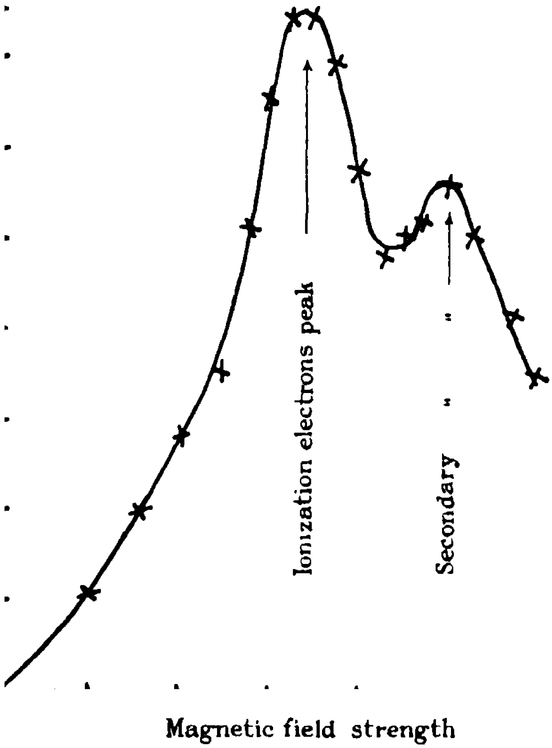


FIG. 3.—Energy of the Hg^+ ions = 1275 volts.

Results.

The curve in fig. 3 shows the variation of the electron current to the collector N with the magnetic field strength H, the Hg vapour being ionized by Hg positive ions of 1275 volts energy. During an experiment the primary ion current and the accelerating potential on the ions were found to remain quite steady. The curve shows two peaks and the magnitudes of the magnetic fields at which these occur are, very approximately, in the ratio of $1 : \sqrt{2}$.

The strength of the magnetic field corresponding to the first peak agrees within the limits of experimental error with the theoretical value given by :—

$$H = \sqrt{\frac{m}{e}} \cdot \frac{\sqrt{2V}}{r},$$

where $V = 40$ volts and $r = 3 \cdot 02$ cm.

Similarly, remembering that "r" remains the same and $V = 80$ volts, this relation holds also for the second peak.

The strength of the magnetic field was determined approximately by observation of the time of oscillation of a small suspended magnet.

The evidence is thus conclusive that the first peak represents electron current coming from the centre of the plates, while the second peak corresponds to electrons from the bottom plate. The large spread of both peaks is probably instrumental.

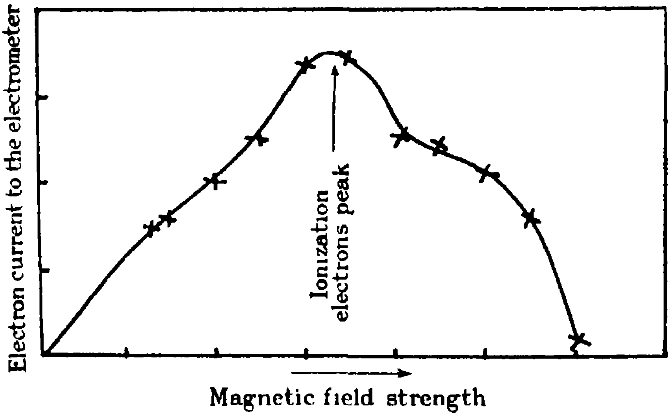


FIG. 4.—Energy of Hg^+ ions = 700 volts.

It has been shown in a previous paper* that the efficiency of emission of secondary electrons from undegassed metal surfaces bombarded by Hg positive ions falls to a very low value for energies below 1000 volts. It would therefore be expected that if a curve be taken with ions of energy less than 1000 volts, the secondary electron peak should decrease relative to the ionization peak.

Fig. 4 shows the results obtained with Hg^+ ions of 700 volts velocity. It can be seen that the second peak has almost disappeared. This supports the

---

[*] Chaudhri, 'Proc. Camb. Phil. Soc.,' vol. 28, p. 349 (1932).

hypothesis that this peak is due mainly to the liberation of secondary electrons from the lower plate by the scattered positive ions.

A curve taken with Hg ions of 1600 volts energy was practically identical with that given in fig. 3, except that the ordinates were increased.

In fig. 5 are given the curves obtained for ionization of mercury vapour by K$^+$ ions of 1600 volts energy. The two curves were taken with different currents and show proportionality between the collector current and the primary beam. Fig. 6 represents the curve obtained with K$^+$ ions of 2400 volts energy. These curves are similar to those found for Hg$^+$ ions. There are, however, some important differences. For instance, the ionization electron peak in the K$^+$ ion curve is smaller than the secondary electron peak, which is the reverse of that found for Hg$^+$ ions. The potassium curve is also

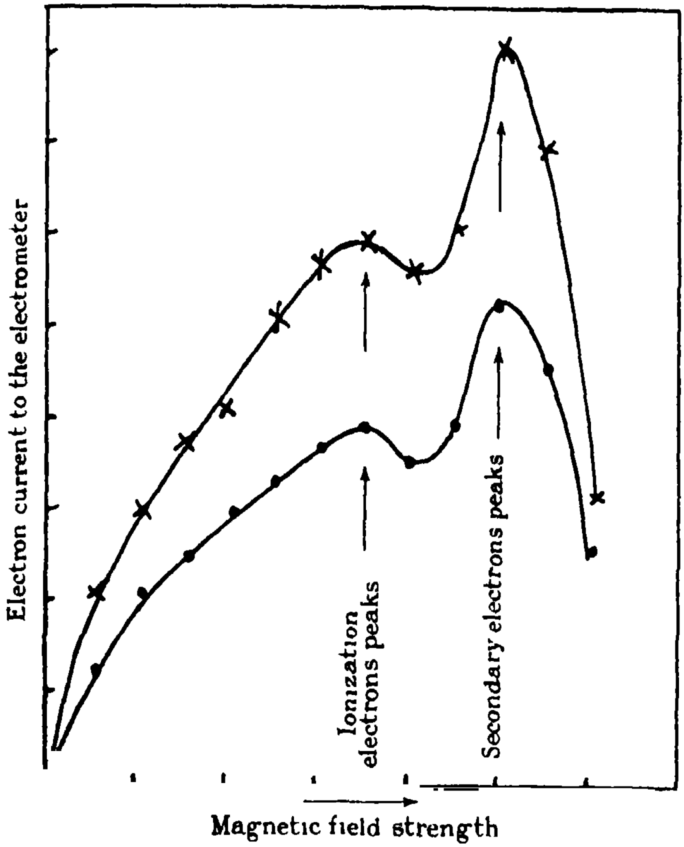


FIG. 5.—Energy of K$^+$ ions = 1600 volts.

more concave to the H-axis than the Hg curve; possibly this is caused by the greater spreading of the K^+ ion beam.

It is possible to calculate relative ionization efficiencies from the curves by rough estimates of the area enclosed beneath the curve drawn symmetrically about the first peak, after converting from i/H to i/H^2 curves. These preliminary estimates are given in Table I.

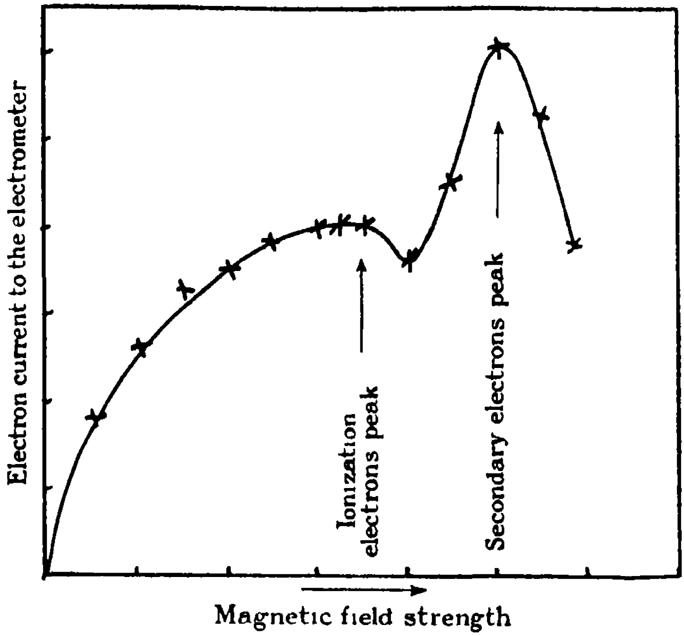


FIG. 6.—Energy of K+ ions = 2400 volts.

Table I.—Ionization of Mercury Vapour at Room Temperaure.

Kind of ions.	Energy of ions.	Relative efficiency of ionization (arbitrary).
	volts	
Hg^+	1275	1·0
K^+	1600	0 76
K^+	2400	1·53

Some tests have been made to show that the ionization produced in Hg vapour was due to positive ions and not to other agencies. With Hg^+ ions it might be suspected that the effect could be due to radiations coming from the arc chamber. It was found, however, that with the arc at full strength there was no electron current to the collector if no accelerating potential was applied to the Langmuir probe. This shows that the radiations from the discharge were not effective. Similarly it was found that the radiations from the hot tungsten plate in the potassium experiment were not effective.

If a side tube attached to the ionization chamber was dipped into liquid air the ionization electron current was found to decrease, as would be expected, on account of the diminished vapour pressure of Hg. The current showed a further decrease if an additional pad of cotton wool, dipped in liquid air, was placed on the glass above the plates. The current, however, rose to its initial values after the pad was removed.

Discussion.

Several theories have been put forward to explain the ionization of gases by positive ions. According to that due to Sir J. J. Thomson,* a Hg^+ ion would require energy of about one million volts in order to ionize a mercury atom. The present experiment shows, however, that ionization takes place at as low an energy of the ion as 700 volts.

Zwicky† has recently suggested that the ionization by ions of moderate energies is a secondary result of a momentum transfer from the impinging ion to the target atom. Since the momentum change is greatest when the two particles are of equal masses it would be expected that the efficiency of ionization would be higher for ions having mass equal to that of the atom, than for ions of different mass but the same kinetic energy.

This is in agreement with the experimental results, given in Table I, which show that Hg^+ ions accelerated by 1275 volts ionize mercury vapour more efficiently than K^+ ions of 1600 volts energy.

The secondary electron emission produced from undegassed metal surfaces by the impact of K^+ ions is of the same order as for Hg^+ ions. The experimental potassium curves show that the secondary electron peak is greater than the ionization electron peak. This is reversed in the mercury curve. This again supports the above conclusion that the ionization of mercury vapour is greater for Hg^+ ions than for K^+ ions of the same kinetic energy.

It is worth mentioning that the secondary electrons coming from the bottom plate are not effective in producing ionization between the plates, the distance between which is quite small compared to their mean free path in mercury vapour at a pressure of 10^{-3} mm. Hg. If they did ionize at all the peak would have occurred at a low magnetic field corresponding to maximum ionization near the upper plate, but the experimental curves show that the first peak corresponds to electrons coming from the centre of the plates.

---

* 'Phil. Mag.,' vol. 48, p. 1 (1924).

† 'Proc. Nat. Acad. Sci. Wash.,' vol. 18, p. 314 (1932).

It was not possible to determine a lower limit for the energy of the ions at which ionization begins to be apparent. With a small accelerating potential on the probe K very few positive ions penetrated the canal D and the resulting current could not be recorded by the galvanometer.

It is my pleasure to record my thanks to Professor Lord Rutherford for his encouragement and interest and Dr. Chadwick for his continued advice. I am indebted to Dr. M. L. Oliphant for his assistance throughout.

Summary.

A method is described by which the ionization of Hg vapour by Hg^+ and K^+ ions of moderate energies has been investigated. This method is shown to be free from many of the defects which existed in the methods of previous workers. It has been shown that Hg^+ ions ionize with greater efficiency than K^+ ions of the same kinetic energy.

It is hoped to continue the experiments and extend their range.

The Structure of Magnesium, Zinc and Aluminium Films.

By G. I. Finch and A. G. Quarrell, Imperial College of Science and Technology.

(Communicated by W. A Bone, F.R S —Received February 25, 1933.)

[Plates 4, 5.]

Introduction.

In connection with investigations now in progress in these laboratories, on the cathodic* and heterogeneous catalytic† combustion of gases and on the electrical condition of metals‡, the need has arisen for a knowledge of the nature of the structure of certain metal films, whether obtained by cathodic sputtering or by thermal evaporation and subsequent condensation.

It has previously been found by spectrographic means that metal sputtered into the cathode zone of a cold-cathode high-tension arc appears therein, at least in part, in the atomic state ; but the results do not enable it to be inferred that aggregates of the atoms are not likewise present. A further important question arising out of the results of our investigations required for its solution the preparation of oxygen-free metal surfaces and their subsequent examination in the absence of oxygen.

The present investigation was therefore undertaken with the main object of developing a technique suitable for the preparation and structural examination of oxygen-free surfaces of metals, both readily oxidizable and otherwise, and of determining whether the crystalline structure, if any, of such surface films was due to the deposition of crystals as such, or to their formation on the receiver subsequent to the arrival of discrete atoms.

Experimental.

The most suitable method at present available for the examination of the crystal structure of thin films of the order of 10^{-6} cm. or less in thickness is by electron diffraction.§ Furthermore, Wierl|| has shown that diffraction also

---

* ' Proc. Roy. Soc.,' A, vol. 133, p. 173 (1931), and seven earlier parts referred to therein.

† ' J. Chem. Soc.,' p. 1544 (1932) and a communication recently submitted to the Society for publication in the ' Proceedings.'

‡ ' Proc. Roy. Soc.,' A., vol. 132, p. 192 (1931) and three earlier parts referred to therein.

§ Thomson and Fraser, ' Proc. Roy. Soc.,' A, vol. 128, p. 41 (1930).

|| ' Phys. Z.,' vol. 31, p. 1028 (1930).

occurs when a jet of molecular vapour is traversed by a suitable electron beam, whereas an atomic gas merely gives rise to electronic scattering.

In view of the necessity for preparing and keeping the films out of all possible contact with oxygen until after the completion of their examination, they were prepared by evaporation *in vacuo* in an electron diffraction camera. In the course of the experiments to be described below, oxygen-free films of magnesium, zinc and aluminium were obtained and it was found, *inter alia*, that the crystalline structure of such films was due to the aggregation of atoms which took place on, or shortly after, their arrival at the receiver upon which the films were deposited. We have also been able to observe remarkable cases of crystal distortion, due to pseudomorphic effects.

The Electron Diffraction Camera.—The essential requirements to be fulfilled by the diffraction camera were as follows: (i) the apparatus should permit of the preparation *in vacuo* of films of readily oxidizable metals and their subsequent electron diffraction examination either *in vacuo* or in the absence of oxidizing or otherwise reacting gases, (ii) the electron beam falling upon the film, should be sufficiently diffuse (*a*) to permit of a rapid examination of a large area of film and (*b*) to minimize the possibility of structural changes due to any thermal or other effects resulting from an excessive intensity of the beam where it impinged upon the specimen, without corresponding loss of definition or intensity at the photographic plate itself; (iii) provision was required to enable a search to be made for possible crystalline structure or aggregation of the metal vapour particles, prior to their deposition.

The electron diffraction camera which fulfilled these requirements is shown diagrammatically in fig 1 The cathode chamber assembly consisted of the glass vessel, A, the aluminium cathode, B, and the brass ring, C, the lower surface of which was lapped dead flat to make a vacuum-tight joint when suitably lubricated with low-vapour-pressure grease. Vacuum-tight joints between these components were made with " Picein " wax.

The water-cooled, cast-aluminium anode block was provided with the main evacuation port, D, which delivered directly into a "Leyboldt" 3-stage mercury-vapour pump. The cathode chamber was evacuated through the tube, E, the effective diameter of which could be varied by means of a valve, F, operated through a ground-glass joint. Gas was supplied to the cathode chamber through the leak port, G, connected to the leak system to be described below. A Swedish-iron " Morse " taper, situated in the vertical axis of the apparatus, carried the anode diaphragm consisting of a 3 mm. bore copper tube spun over at its upper end to 0·25 mm. diameter. The main tube, of

No. 16 S.W.G. brass, was fitted with side ports, P_1, P_2, P_3, and P_4, and was joined to the anode and camera blocks with picein wax. P_1 carried the specimen chamber assembly to be described below; P_2 served as an observation window, and P_3 was fitted with a lapped brass block carrying the metal evaporating assembly. This consisted essentially of a tungsten filament spot-welded across the two electrodes, e_1 and e_2, one of which was in electrical connection with the main body of the apparatus, the other being led through a glass insulator

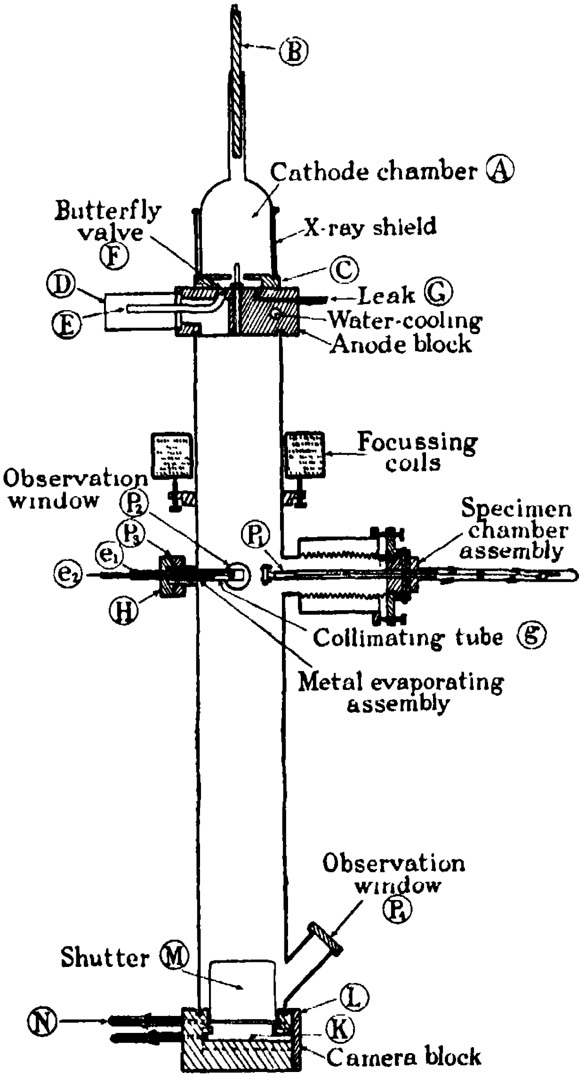


Fig. 1.—Electron diffraction apparatus.

waxed into the block, H. A brass tube, *g*, 8 mm. bore and surrounding the filament served to direct the metal vapour in the form of an approximately parallel beam towards the specimen plate. The port, P_4, was closed by a plate-glass observation window. A focussing coil of about 3000 ampere turns maximum, the axis of which could be tilted relative to that of the apparatus, was situated as shown in fig. 1. The slot, K, in the gunmetal camera block, carried the 6 × 9 cm. photographic plate and was closed by the lapped brass slide, L. The aluminium plate shutter, M, was backed underneath with lead sheet 3 mm. thick, the upper surface being coated with zinc sulphide. The screen could be raised, and the photographic plate thus exposed, by rotating the ground glass joint, N.

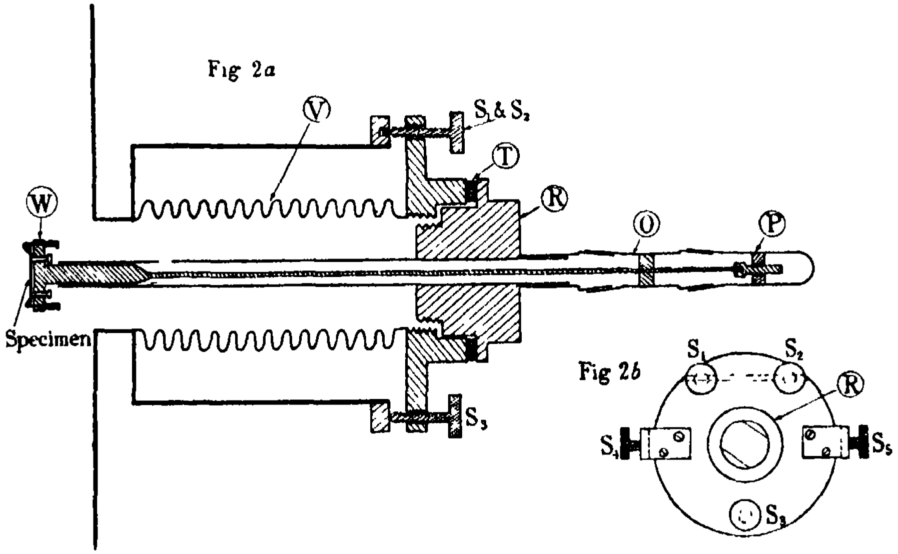


Fig. 2.—Specimen chamber apparatus.

The specimen chamber assembly is shown in fig. 2 in sufficient detail to render clear the manner whereby the following independent adjustments of the specimen, *i.e.*, the metal film under examination, could be effected: (i) inclination of the plane of the specimen to the electron beam—by means of the levelling screws, S_1, S_2 and S_3, the requisite degree of flexibility being provided for by the metallic bellows, V; (ii) translation into the beam—by turning O, while P was held in such a manner as to prevent its rotation; (iii) rotation in the azimuthal plane—by turning O and P together and, finally, (iv) lateral motion in the specimen plane at right angles to the beam—by means of the screws, S_4 and S_5, which also served as a locking clamp.

The actual specimen carrier could be removed from the specimen chamber by unscrewing R, the joint being otherwise rendered vacuum-tight by the fine-grade "Hallite" washer, T. The specimen, consisting of a film of the metal under examination deposited on a 2 cm. diameter flat, polished quartz plate, was clamped on to the disc, W, which was provided with three levelling screws. These enabled the specimen plane of the quartz plate to be adjusted until at right angles to the axis of rotation of W.

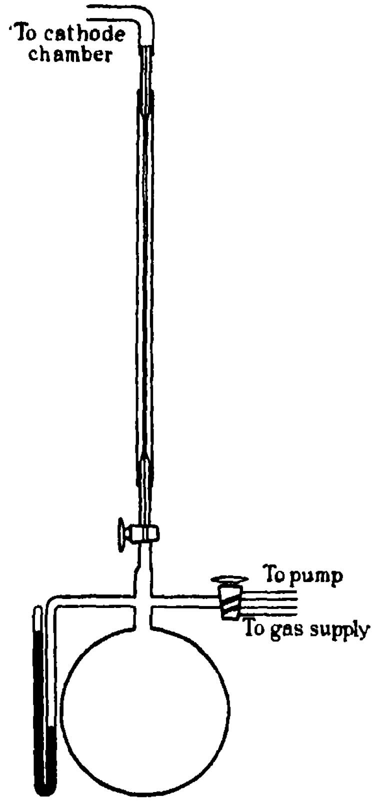


FIG. 3.—Leak system.

The cathode chamber leak, fig. 3, consisted in the main of a fine-bore capillary drawn out in a flame from a short section of 1 mm. bore thick-walled glass tubing and protected by encasing in a suitable glass tube. One end of the capillary leak was waxed into the leak tube of the anode block, the other was in connection with a gas reservoir, a manometer and with suitable vacuum and gas-supply services, as shown in fig. 3. The rate of leak was determined by the pressure in the reservoir, the 1·5 litre volume of which sufficed to ensure a virtually constant rate of leak over long periods.

The High Tension Supply. — Constant current high-tension D.C. was furnished to the cathode as shown in fig. 4 which is self-explanatory, except as regards the water-cooled saturated diode incorporated in the circuit. The functions fulfilled by this valve may be summed up briefly as follows: (i) by controlling the filament temperature the discharge tube current could be varied over a wide range independently of the cathode chamber gas pressure; (ii) the current passed by the diode when saturated was independent of the voltage drop across the smoothing condenser; thus perfect smoothing and hence also a monochromatic electron beam could be obtained with a minimum outlay in smoothing equipment; finally, (iii) since the maximum discharge currents employed were of the order of only 0·5 mA. or less, the use of the saturated diode as a current-limiting device greatly increased the safety of operation. As protection against X-rays, the saturated diode was also almost completely encased in a 3 mm. thick lead sheet chamber.

Preparation of the Films.—Magnesium and aluminium films were obtained by evaporation of short lengths of wire wrapped round the tungsten filament of the metal evaporating assembly. For zinc, the filament was first dipped into the molten metal. Evaporation was carried out under vacuum conditions such that the discharge tube was hard to 90 kV. or more, the whole diffraction apparatus having been previously flushed out with oxygen-free hydrogen. The distance between the heating filament and the specimen plate was of the order of 2 cm. The average time required for the deposition of a suitable film was between 10 and 15 seconds. For reasons which will be made clear below, a platinum film was cathodically sputtered in argon on to the quartz plate before its insertion into the diffraction apparatus. Thus, the films prepared *in vacuo* were deposited on a substrate of sputtered platinum.

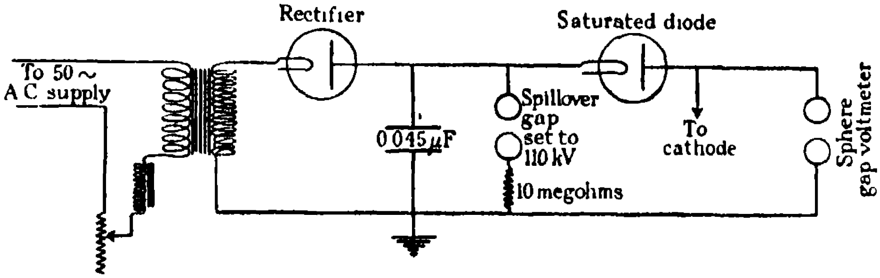


FIG. 4.—High tension current.

The Electron Diffraction Examination.—Prior to depositing a film, the specimen plate was adjusted in a suitably focussed and biassed cathode beam of the required intensity and velocity, until the normal diffraction pattern due to the face-centred cubic structure of the platinum film was plainly visible on the fluorescent screen. It will be realized from fig. 1 that the relative dispositions of the focussing coil, specimen and anode diaphragm were such as to ensure the specimen being swept by the electron beam near the region of its maximum width, the photographic plate lying in the focal plane of the coil. The cathode chamber leak was then closed and the metal under examination evaporated as soon as the diffraction apparatus was sufficiently evacuated. Immediately after deposition of a film, the tungsten filament heating current was interrupted, and hydrogen admitted through the leak at a suitable rate. The discharge tube current was then switched on, and the photographic plate immediately exposed. In general the time elapsing between preparation of a film and the photographic recording of the corresponding electron diffraction pattern was less than half a minute. The discharge tube current and voltage were usually between 0·1 and 0·5 mA. and 30 and 50 kV. respectively. The time of exposure was generally about 1/5 second.

A cathodically sputtered substrate of platinum was employed for the reception of the films in order to examine the effect, if any, of the substrate on the deposited film. Furthermore, it was desirable to be able to adjust beam and specimen plate into their relative positions for diffraction prior to the deposition of the evaporated film.

Calibration of the Diffraction Apparatus.—Magnetic focussing and biassing of the beam to one edge of the photographic plate introduced two sources of error, the one due to non-uniformity of the magnetic field in planes at right angles to, and up to the limits of their intersections with, the cones of diffracted electrons; and the other due to the beam axis not being exactly perpendicular to the photographic plate. In order to determine the magnitude of such errors a series of twelve diffraction patterns were obtained with sputtered platinum films over a range of accelerating voltages of 15 to 65 kV. In the case of each photograph a value for the side of unit cube of platinum was obtained from the radius of each ring. It was found that no systematic error occurred with increasing ring radius, from which it may be concluded that the magnetic field was uniform in the sense outlined above. Furthermore, the average values for the side of unit cube as determined from each plate agreed with X-ray values to within the limit of experimental error amounting to not more than 2%, this error being due mainly to the difficulty of determining the accelerating potential with the sphere-gap meter to an accuracy greater than between 3 and 4%. The average value for the side of unit cube of platinum cathodically deposited on fused quartz was 3·89 A.

Analysis of the Diffraction Patterns.—The patterns were analysed as follows. The value of d/N for each ring was calculated from the relationship $d/\mathrm{N} = \lambda\mathrm{L}/\mathrm{R}$, where d is the spacing of the Bragg planes, N is the order of the reflection, λ the wave-length associated with the electrons, L the effective length of the camera from specimen to plate and R is the radius of the ring. The values of d/N, plotted on a logarithmic scale and fitted to nomograms due to Hull and Davey,* gave the type of unit cell together with the order of its axial ratio and the indices of the Bragg planes. For a triangular close-packing lattice (hexagonal system), the relationship

$$\mathrm{N}^2/d^2 = 4(h_1^2 + h_2^2 + h_1h_2)/3a^2 + h_3^2/c^2a^2, \qquad (1)$$

---

* 'Phys. Rev,' vol. 17, p. 549 (1921).

where the h's are the Millerian indices, a is the basal axis and c the axial ratio. a can be calculated for those planes whose third index $h = 0$. The average of all values of a obtainable in this manner from the pattern, on substitution in equation (1), gave a value of c for those rings for which $h \neq 0$, the resulting average value of c serving to give a for these planes. For crystals of tetragonal structure,

$$\mathrm{N}^2/d^2 = (h_1^2 + h_2^2)/a^2 + h_3^2/c^2a^2, \qquad (2$$

the cubic crystal being a special type for which $c = 1$.

The Results.

The effective length of the camera from specimen to photographic plate was 47·5 cm. throughout.

Table I.—Magnesium deposited on face-centred cubic platinum.—Specimen No. 1.

Photograph 44. Exposure, 1/5 second at 37 kV. $\lambda = 0{\cdot}06236$ A.

The spacings are fitted to the nomograms for a triangular, close-packing lattice (hexagonal system).

Radius in cm.	d/N in A.	Plane.	a in A.	c.
1·04	2·85	(1.0.0)	3·29	—
1·17	2·53	(1.0.1)	3·35	1·68
1·35	2·19	(1.2.0)1/2	3·34	—
1·51	1·96	(1.0.2)	3·26	1·65
1·83	1·62	(1.1.0)	3·24	—
1·95	1·52	(1.0.3)	3·29	1·64
2·07	1·43	(1.0.0)2	3·31	—
2·16	1·37	(2.0.1)	3·28	1·56
2·50	1·19	(1.0.4)	3·14	1·59
2·67	1·11	(2.0.3)	3·28	1·62
2·80	1·06	(1.2.1)	3·30	1·78
2·88	1·03	(1.1.4)	3·21	1·61
3·03	0·98	(1.0.2)2	3·25	1·64
3·13	0·95	(1.0.0)3	3·29	—
3·21	0·93	(1.2.3)	3·30	1·64
3·29	0·90	(3.0.2)	3·30	1·71
3·54	0·84	(1.2.4)	3·25	1·63
3·60	0·83	(1.1.0)2	3·30	—
3·79	0·78	(1.1.6)	3·29	1·63

Mean value of $a = 3{\cdot}28$ A. **Mean value of $c = 1{\cdot}64$.**

Table II.—Specimen No. 2, fig. 5, Plate 4.

Photograph 72. Exposure, 1/5 second at 43 kV. λ = 0·05763 A.

Triangular, close-packing lattice (hexagonal system).

Radius in cm.	*d*/N in A.	Plane.	*a* in A.	*c*.
0·96	2·85	(1.0 0)	3·29	—
1·10	2·49	(1 0.1)	3·26	1·71
1·43	1·92	(1.0.2)	3·27	1·64
1·70	1·61	(1.1.0)	3·22	—
1·84	1·49	(1 0.3)	3·28	1·64
1·97	1·39	(1.0 0)2	3·21	—
2·05	1·34	(2.0 1)	3·20	1·45
2·24	1·22	(1 0 1)2	3·22	1·58
2·54	1·08	(2 0 3)	3·21	1·59
2·62	1·05	(1.2.0)	3·20	—
2·67	1·03	(1.2.1)	3·20	1·41
2·70	1·02	(1.1.4)	3·26	1·62
2·84	0·97	(1 0.5)	3·22	1·60
2·98	0·92	(1 0 0)3	3·19	—
3·04	0·90	(1.2.3)	3·23	1·62
3·06	0·89	(0 0.1)6	—	1·67
3·10	0·88	(3.0.2)	3·22	1·78
3·20	0·86	—	—	—
3·63	0·76	(1.1.6)	3·18	1·60

Mean value of *a* = 3·23 A. Mean value of *c* = 1·62.
Mean value of *a* from both photographs = 3.25 A.
Mean value of *c* from both photographs = 1·63.
The corresponding X-ray values* are 3·22 A. and 1·624.

---

* All X-ray values quoted in this communication have been taken from the 'Handbuch d. Physik,' vol. 24, pp. 329–330 (1927).

Table III.—Magnesium Oxide formed by oxidation of magnesium in air.—Specimen No. 2, fig. 6, Plate 4.

Photograph 75. Exposure, 1/5 second at 42 kV. λ = 0·05839 A.

Simple triangular lattice (hexagonal system).

Radius in cm.	*d*/N in A.	Plane.	*a* in A.	*c*.
1·07	2·60	(1.0.0)	2·99	—
1·23	2·25	(1.0.1)	2·98	1·55
1·75	1·59	(0.0.1)3	3·05	1·55
1·95	1·42	(1.1.1)	2·99	1·51
2·15	1·29	(1.0.0)2	2·98	—
2·46	1·13	(1.0.1)2	2·98	1·55
2·76	1·01	(2.0.3)	3·02	1·61
2·83	0·98	(1.2.0)	2·99	—
3·02	0·92	(1.1.4)	2·99	1·56
3·44	0·81	(3.0.2)	2·98	1·53
3·72	0·75	(1.1.0)2	2·98	—

Mean value of *a* = 2·99 A. Mean value of *c* = 1·55.

Table IV.—Zinc deposited on face-centred cubic platinum.—
Specimen No. 1, fig. 7, Plate 4.

Photograph 54. Exposure, 1/5 second at 42 kV. $\lambda = 0{\cdot}05839$ A.
Triangular, close-packing lattice (hexagonal system).

Radius in cm.	d/N in A.	Plane.	a in A.	c.
1·23	2·26	(1.0 0)	2·60	—
1·36	2·04	(1.0.1)	2·57	1·74
1·61	1·72	(1.0.2)	2·65	2·00
2·04	1·36	(1.0.3)	2·60	1·93
2·11	1·31	(1.1.0)	2·62	—
2·17	1·28	(0.0.1)4	2·59	1·94
2·35	1·18	(1.1.2)	2·65	2·03
2·40	1·16	(1.0.0)2	2·67	—
2·49	1·11	(2.0.1) (1.0.4)	2·60 2·63	2·03 1·95
2·65	1·04	(1.0.1)2	2·64	2·02
2·92	0·95	(2.0.3)	2·66	1·97
3·02	0·92	(1.1.4)	2·62	1·95
3·26	0·85	(0.0.1)6 (1.0.2)2 (1.2.0)	2·59 2·60 2·60	1·95 1·95 —
3·38	0·82	(1.2.2)	2·64	2·06
3·61	0·77	(1.2.3)	2·62	1·95
3·80	0·73	(3.0.2)	2·63	1·98
3·91	0·71	(1.2.4)	2·60	1·90
4·00	0·69	(1.0.7)	2·59	1·94
4·23	0·66	(1.2.5)	2·61	1·92
		Also seven unidentified rings of radius $>$ 4·3 cm.		

Mean value of $a = 2{\cdot}62$ A. Mean value of $c = 1{\cdot}97$.

In addition there were several rings within this range which did not agree with this structure; these rings, however, fitted the nomogram at the point for which $c = 2{\cdot}58$. At the same point, some of the above rings also fitted and these are included in the analysis given in Table VI and denoted by asterisks. Seven rings with d/N between 0·62 and 0·54 A. could not be identified and these are omitted. Among the additional rings there was no ring from which a could be obtained directly, and the value of c obtained from the nomogram was therefore assumed and the values of a calculated in the usual manner.

Magnesium, Zinc and Aluminium Vapours.—No diffraction patterns were obtained with the electron beam traversing a stream of vapour of any of the above metals. On the other hand, intense blackening, which faded off towards the edges of the photographic plate, afforded ample evidence of profuse scattering of electrons by the vapour. In these experiments the times of evaporation were varied between approximately 5 and 15 seconds, and in each exposure

Table V.—Zinc-Zinc Oxide formed by short exposure of zinc, Specimen No. 1, to air.—A pattern due to both Zn and ZnO was obtained. The following are the rings due to zinc, fig. 8, Plate 4.

Photograph 55. Exposure, 1/5 second at 38 kV. $\lambda = 0\cdot06150$ A.

Triangular, close-packing lattice (hexagonal system).

Radius in cm.	d/N in A.	Plane.	a in A.	c.
1·14	2·56	(0.0.1)2	2·72	1·91
1·26	2·32	(1.0.0)	2·68	—
1·38	2·12	(1.0.1)	2·68	1·95
1·68	1·74	(1.0.2)	2·70	1·96
2·11	1·39	(1.0.3)	2·68	1·93
2·17	1·35	(1.1.0)	2·70	—
2·25	1·30	(0.0.1)4	2·69	1·94
2·45	1·19	(1.1.2)	2·69	1·95
2·50	1·17	(1.0.0)2	2·68	—
2·58	1·13	(2.0.1) (1.0.4)	2·68 2·69	1·93 1·94
2·75	1·06	(1.0.1)2	2·69	1·97
3·06	0·96	(1.0.5) (2.0.3)	2·71 2·67	1·95 1·88
3·16	0·93	(1.1.4)	2·66	1·91
3·35	0·87	(1.2.0)	2·67	—
3·40	0·86	(1.0.2)2 (0.0.1)6	2·67 2·67	1·91 1·92
3·55	0·84	(1.2.2)	2·66	1·78
3·66	0·80	(1.0.6)	2·65	1·91
3·72	0·79	(1.2.3)	2·69	1·97
3·76	0·78	(2.0.5) (1.0.0)3	2·70 2·69	1·95 1·95
3·96	0·74	(3.0.2)	2·67	1·83
4·05	0·72	(1.0.7) (1.2.4)	2·74 2·66	1·98 1·89
4·37	0·67	(1.1.0)2	2·68	—

Mean value of $a = 2\cdot68$ A. Mean value of $c = 1\cdot93$.
Mean value of a for both photographs $= 2\cdot65$ A.
Mean value of c for both photographs $= 1\cdot95$.
Corresponding X-ray values 2·67 A. and 1·86.

the intensity of the beam was suitably reduced to permit of the discharge being switched on shortly after, and subsequently maintained throughout, deposition without leading to over exposure. Furthermore, in different experiments the electron beam was made to traverse the vapour jet at various points between the collimating tube surrounding the evaporating filament and the specimen plate without, however, being allowed to impinge upon either.

The Appearance of the Films.—The appearance of the metal films was observed through the window, P_2, the necessary illumination being provided by the tungsten evaporating filament. The freshly deposited films had perfectly clean, bright metallic mirror surfaces. On admission of calcium

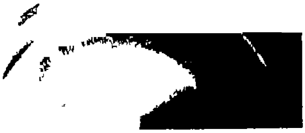


Fig. 5.—Magnesium

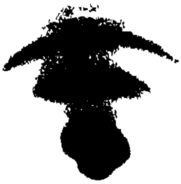


Fig. 6.—Magnesium oxide

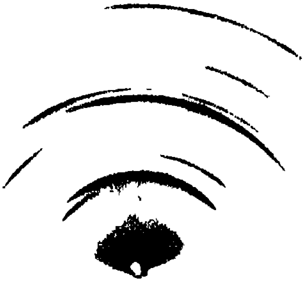


Fig. 7.—Zinc

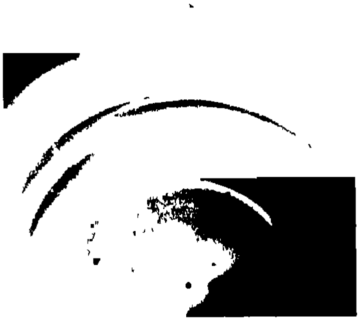


Fig. 8.—Zinc oxide and zinc

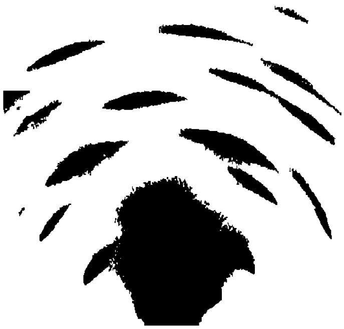


Fig. 9.—Aluminium (tetragonal)

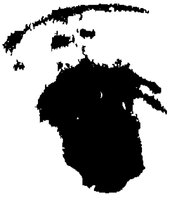


Fig. 10.—Aluminium (cubic)

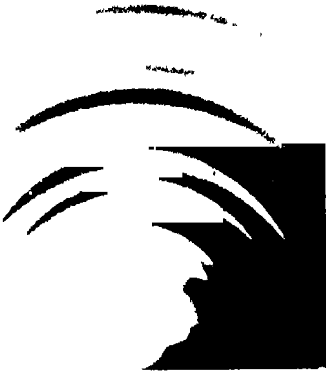


Fig. 11.—Platinum.

Fig. 12.—Platinum

Table VI.—Rings due to zinc oxide.

Triangular, close-packing lattice (hexagonal system).

Radius in cm	d/N in A.	Plane.	a in A.
1·36	2·15	(1 0 1)	2·62
1·73	1·69	(0 0.1)4	2·60
1·80	1·62	(1 0 3)	2·66
2·39	1·22	(1 1.2)	2·62
*2·50	1·17	(1.0 5)	2·64
*2·58	1·13	(2 0 1) (0.0.1)6	2·70 2·63
3·28	0·89	(1 0.7)	2·63
*3·35	0·87	(2.0 5)	2·63
*3·40	0·86	(1 2 0)	2·63
*3·66	0·80	(1 2 3)	2·60
*3·72	0·79	(1.0.8)	2·60
*3·96	0·74	(2.0 7)	2·62
4·29	0·68	(0 0.1)10 (1.0 4)2	2·64 2·60
4·43	0·66	(1.1 0)2	2·64
4·51	0·65	(1.2.7)	2·65
4·55	0·64	(1.0.10)	2·60

Mean value of a = 2·63 A.

Mean value of c as given by graph = 2·58 (N B.—This value was obtained from the rings not marked with an asterisk.)

Corresponding X-ray values for zinc oxide are : a = 3·22 A. and c = 1·61.

chloride-dried air to the diffraction camera, however, the zinc and magnesium films rapidly lost their brilliant lustre, the surfaces becoming matt and frosted. Under similar circumstances, aluminium films, whilst retaining to some extent their metallic appearance, lost somewhat in lustre and took on a bluish tint.

Discussion.

The results set forth above prove that, under the conditions of our experiments, magnesium, zinc and aluminium vapours are monatomic, and that crystallization occurs only on or after condensation and deposition on the specimen plate. In this respect these metal vapours behave whilst in transit between source and receiver as do cathodically sputtered metals.* In view of these facts it seemed reasonable to suppose that the crystalline structure of the deposited films might well be influenced by the nature of the substrate. Such an effect was accordingly looked for in the case of the metals now under discussion.

It was shown in the course of calibrating the diffraction apparatus that the side of unit face-centred cube of platinum deposited on amorphous quartz

---

* 'Proc. Roy. Soc.,' A, vol. 124, p. 303 (1929); vol. 129, p. 314 (1930).

Table VII.—Aluminium deposited as a thin layer on face-centred cubic platinum. -Specimen No. 1, fig. 9, Plate 5.

Photograph 46. Exposure, 1/5 second at 45 kV. $\lambda = 0{\cdot}05635$ A.

Face-centred tetragonal lattice.

Radius in cm.	*d*/N in A.	Plane.	*a* in A.	*c*.
1·17	2·20	(1.1 1)	3·92	1·07
1·33	2·01	(0 0.1)2	3·90	1·04
1·95	1·37	(1.1.0)2	3·89	—
2·25	1·19	(1.1.3)	3·86	1·02
2·31	1·16	(1.3 1)	3·84	0·92
2·36	1·14	(1.1.1)2	3·81	1·04
2·75	0·98	(1.0.0)4	3·90	—
2·98	0·90	(1.3 3)	3·87	1·02
3·05	0·88	(1.0 2)2	3·84	1·02
3·12	0·86	(1.2 0)2	3·84	—
3·34	0·80	(1.1 2)2	3·86	1·02
3·41	0·79	(1.2 1)2	3·83	0·96
3·54	0·76	(1.1.1)3 (1.1.5)	3·91 3·83	1·06 1·02
3·61	0·74	(1 5.1)	3·82	0·99
3·90	0·69	(1.1.0)4	3·89	—
3·97	0·67	(1.3 5)	3·89	1·04
4·08	0·66	(1.5.3) (1.2.2)2	3·88 3·90	1·04 1·04
4·15	0·65	(3 5 1) (1 0 0)6 (2 2 1)2	3·84 3·88 3·88	1·06 — 1·07
4·35	0·62	(1.3.0)2	3·90	—
4·41	0·61	(3.3 5) (1.1.3)2	3·94 3·94	1·05 1·02
4·50	0·60	(3.5.3)	3·89	1·04
4·60	0·58	(1 3 1)2	3·86	0·98
4·89	0·55	(2.0 3)2	3·89	1·03
5·00	0·54	(2.3.0)2	3·88	—
5·29	0·51	(1.7.3)	3·89	1·04

Mean value of $a = 3{\cdot}90$ A. Mean value of $c = 1{\cdot}03$.

is 3·89 A., the normal X-ray value being 3·91 A. The normal structure of aluminium as has been revealed by X-rays, is likewise that of a face-centred cubic lattice, the side of unit cube being 4·05 A. We have, however, been able to obtain a thin layer of aluminium deposited on platinum which possessed a face-centred tetragonal structure, the axial ratio being 1·03 (photograph 46). On the other hand, a sufficiently thick layer of aluminium gave a diffraction pattern corresponding to a face-centred cubic lattice, the side of unit cube being 4·01 A. (photograph 77). Thus the ratio, the side of unit cube of aluminium to that of platinum, as determined above by electron diffraction is $1{\cdot}03_0$ and by X-rays $1{\cdot}03_8$; values which approximate closely to that of the axial ratio of unit cell of aluminium deposited as a thin film on

Table VIII.—Aluminium deposited as a thick layer on face-centred cubic platinum.—Specimen No. 2, fig. 10, Plate 5.

Photograph 77. Exposure, 1/5 second at 51 kV. $\lambda = 0{\cdot}05277$ A.

Face-centred cubic lattice.

Radius in cm.	d/N in A.	Plane.	a in A.
1·06	2·36	(1.1.1)	4·10
1·24	2·02	(1 0.0)2	4·04
1·76	1·42	(1.1.0)2	4·03
2·10	1·19	(1.1.3)	3·96
2·48	1·01	(1.0.0)4	4·04
2·80	0·89	(1.2.0)2	4·00
3·08	0·81	(1.1.2)2	3·99
3·30	0·76	(1.5.1) (1.1.1)3	3·95 3·95
3·74	0·67	(6.0.0) (4 4 2)	4·02 4·02

Mean value of side of unit cube = 4·01 A. Corresponding X-ray value = 4·05 A.

platinum. Finally, the length of the basal axis, a, of the face-centred tetragonal aluminium unit cell, as determined in the experiments outlined above, is 3·90 A., a value nearly equal to that of the side of the unit cube of platinum, whilst the length of the major axis is $3{\cdot}90 \times 1{\cdot}03 = 4{\cdot}02$ A, *i.e.*, approximately the length of the side of the unit cube of aluminium. In view of these facts it is difficult to resist the conclusion that when aluminium vapour is deposited on face-centred cubic platinum, under the conditions of our experiments, the platinum atoms exert attractive forces on the aluminium atoms arriving to form the first plane, thus causing them to take up positions in the lattice similar to those which would have been occupied by platinum atoms had the vapour consisted of this metal. The aluminium atoms of this layer, in their turn, determine the placing of further atoms, causing these on arrival to occupy positions with similar interatomic separations in the next parallel plane, and so on. The distance between such parallel aluminium planes, however, is governed mainly by the normal properties of aluminium and is thus characteristic of this metal, except so far as a drawing together, *i.e.*, a packing closer than normal of the aluminium atoms in the individual basal planes, may tend to increase the spacing of such planes. On the other hand, the fact that the aluminium atoms can be drawn together on the platinum lattice shows that the atoms are not close-packed in the normal aluminium face-centred cube. The effect upon the basal plane spacing of pseudomorphic closer-packing within these planes is therefore not likely to be large. Indeed, our experimental results show that the effect, if any, is so small as to have escaped detection. The fact that a sufficiently thick aluminium layer, even though deposited on face-centred cubic platinum, exhibits the normal structure of aluminium, shows that the effect of the above-outlined pseudomorphic strain upon the surface layers disappears with sufficient film thickness. The aluminium films could easily be rubbed off the substrate of sputtered platinum without apparent injury to the latter. This fact suggests that no actual alloying occurred between the two metals.

No such pseudomorphic phenomena were observed with either magnesium or zinc deposited on platinum, probably owing to the fact that only relatively thick films were examined. The structures obtained for the corresponding *extremely thin oxide layers, produced by exposing the metal films to dry air*, however, are not those which have previously been obtained by X-ray examination of the oxides in bulk. It therefore seems reasonable to suggest that the difference in structure between the oxides in thin layers on the corresponding metals and the metal-free oxides as usually serving for X-ray examination may be accounted for by pseudomorphic effects. Thus, for example, the lengths of the basal axes of zinc and zinc oxide in bulk as determined by X-rays are 2·67 and 3·22 A. respectively, whereas the corresponding electron diffraction values for the metal and for a thin oxide film on the metal are 2·65 and 2·63 A. Moreover, the volume of the unit cell of zinc oxide in bulk is, by X-rays, 46·6 A.^{3}, whilst that of a thin zinc oxide film on zinc has now been found by electron diffraction to be 40·6 A.^{3}. Since the possible experimental error inherent in the above electron diffraction determinations of axial lengths is about 2 per cent., it will be realized that the difference of 6·0 A.^{3} between the two volumes barely exceeds this error. Thus it would appear that the zinc oxide cell, in forming on the metal surface, takes up basal dimensions similar to those of the zinc cell, the contraction thus involved being counterbalanced by a suitable increase in the axial ratio, with the result that the volume of the unit cell tends to remain substantially unchanged.

The crystal structure of the magnesium films has been shown above to be normal, *i e*, that of a close-packed triangular lattice (hexagonal system), whereas that of the oxide film formed thereon is a simple triangular lattice (hexagonal system) and therefore differs from the normal face-centred cubic lattice structure of magnesium oxide in bulk. Whilst these facts afford clear evidence of pseudomorphic effects again coming into play during the formation of the magnesium oxide on the metal surface, the crystal structures of the metal and oxide films differ to such an extent that a quantitative explanation on the lines followed in the two cases previously discussed cannot be put forward at this stage.

The Effect of Heat upon the Crystal Structure.—It will be seen from the diffraction patterns, figs. 5 to 10, that the crystals were sometimes orientated in certain preferred directions, whilst in others, as shown for example by fig. 10, Plate 5, the crystal array was a completely random one. Kirchner* has recently observed a similar phenomenon and has shown that its occurrence is due to thermal effects. In our experiments the sputtered platinum films employed as substrates for the reception of the vaporized metals possessed a completely unorientated structure, as is shown by fig. 11, Plate 5. If the conditions of vaporization were such as to avoid any pronounced heating of the receiver, films were obtained in which the crystals were disposed in random directions. On the other hand, when the receiver became sufficiently heated as a result of radiation from the tungsten evaporating filament, the crystals took up a general alignment in certain preferred directions. It seems reasonable to suppose that, under such conditions, the substrate likewise shared in the orientation, because we have found that the random dispositions of the platinum crystals, giving rise to a diffraction pattern such as is shown in fig 11, changed on heating by radiation from the tungsten filament, to an ordered array, in the sense that the crystals were strongly orientated, as is shown by fig. 12, Plate 5.

We wish to thank Imperial Chemical Industries, Ltd., for a grant which defrayed the cost of building the diffraction apparatus and also part of that of the high tension outfit. One of us (A.G Q.) wishes to thank the Department of Scientific and Industrial Research for a grant which enabled him to devote his whole time to this work.

Summary.

An electron diffraction camera is described in which the specimen is swept by a diffuse beam of electrons, thereby reducing the risk of injury to the specimen. The beam is subsequently focussed electromagnetically.

Oxide-free surfaces of magnesium, zinc, and aluminium on a platinum substrate, also vapours of these metals in transit between source and receiver, were obtained and examined by electron diffraction. The oxides of magnesium and zinc formed on the corresponding metal surfaces have also been examined.

---

* 'Phys. Z.,' vol. 76, p. 576 (1932).

It has been found that (i) the structure of a thin aluminium film deposited on platinum is that of a face-centred tetragonal lattice and thus differs from that of the normal structure which, however, is obtained with sufficient film thickness; (ii) thin oxide layers of magnesium and zinc likewise exhibit abnormal crystal structure; (iii) heating of the receiver causes the crystals in the films hitherto examined to orientate themselves in certain preferred directions; (iv) the metal vapours do not diffract but scatter the electrons.

From these facts it is concluded that (i) the abnormal crystal structures are due to pseudomorphic strain effects whereby the substrate influences the positions taken up by the atoms of the metal or metal oxide layers formed thereon; (ii) such pseudomorphic effects are no longer evident at the surface of sufficiently thick films; (iii) in the case of aluminium the pseudomorphic strain effects are confined to the two dimensions of the basal planes, because no such effect was to be observed in the third dimension; (iv) magnesium, zinc and aluminium vapours are monatomic.

The Catalytic Properties and Structure of Metal Films.—Part I. *Sputtered Platinum.*

By G. I. Finch, C. A. Murison, N. Stuart, and G. P. Thomson, F.R.S., Imperial College of Science and Technology, London.

(Received February 27, 1933.)

[Plates 6–7.]

Introduction (G.I.F. and N.S.).

During an experimental study by Stimson and one of us* of the electrical condition of a platinum sheet, it was found that the surface sometimes exhibited a charge at room temperature after a previous treatment consisting, in the main, of heating in oxygen and hydrogen alternately. It was found later that cathodically sputtered platinum films were not only frequently electrically charged at room temperature, but were also sometimes exceedingly active catalysts of the union of hydrogen and oxygen at room temperature. In the course of these experiments, it soon became evident that the conditions under which the sputtered films were prepared in some manner profoundly affected their catalytic properties, and the following investigation was therefore carried out with the object, *inter alia*, of determining in what manner and to what extent sputtering conditions influenced the catalytic and structural properties of the resulting films.

---

* 'Proc. Roy. Soc.,' A, vol. 124, p. 356 (1929).

The experiments of which an account is now given in the ensuing first part of this communication were begun by two of us (G. I. F. and N. S.) in September, 1929. Both catalytically active and inactive films were obtained, and the sputtering conditions determining their production sufficiently well defined to enable the different types of surface to be prepared at will. It soon became evident, however, that the activity or otherwise of the films was in some manner intimately associated with structural properties which, owing to their submicroscopic nature, defied direct examination by the means then at our disposal.

In May, 1930, Mr. H. M. D. White at our request carried out at University College an X-ray examination by the Debye-Scherrer method of two platinum films which had been sputtered for this purpose on to quartz rods. The results were, however, negative. Accordingly, in March, 1931, it was suggested to Professor G. P. Thomson, F R S , that an electron diffraction examination of these platinum surfaces might possibly disclose that feature of the structure upon which their catalytic properties appeared to depend. A preliminary examination led to such promising results that it was agreed to collaborate in a study of sputtered platinum films on the lines set forth above ; the preparation, chemical and microscopic examination and determination of the catalytic properties to be carried out by two of us (G. I. F. and N. S.), and the sub-microscopic structure of the films to be investigated by the two other authors by electron diffraction.

The experimental results set forth below throw fresh light, not only upon the mechanism of the catalytic combustion of hydrogen by oxygen, but also upon the rôle played by metal atoms in promoting the cathodic combustion of carbonic oxide and of hydrogen.* The investigation is therefore being extended to the further study of platinum and other metal films with the main object of determining the relationship between the catalytic activity and specificity of action of such films and their nature and structure.

---

* 'Proc. Roy. Soc.,' A, vol. 124, p. 303 (1929) ; vol. 125, p. 532 (1929) ; vol. 129, pp. 656 and 672 (1930) ; vol. 133, p. 173 (1931).

Experimental.

Apparatus.—Sputtered platinum films were prepared in the apparatus shown diagrammatically in fig. 1. A wide-mouthed glass vessel closed by a ground-in stopper was fitted with two hook-shaped platinum anodes, A, and a similar but axially disposed cathode hook, K, from which was suspended a length of No. 18 S.W.G platinum wire. The receivers, R, upon which the films were deposited by sputtering, consisted either of a pair of transparent quartz or glass plates, a molybdenum block and a glass plate, or a 1·25 cm. bore glass tube suspended from the anode hooks co-axially with the cathode. The complete sputtering assembly was surrounded by the wide-bore glass tube shield, S, which served to prevent deposition of sputtered metal on the walls of the main vessel. The large ground-in joint was made vacuum-tight

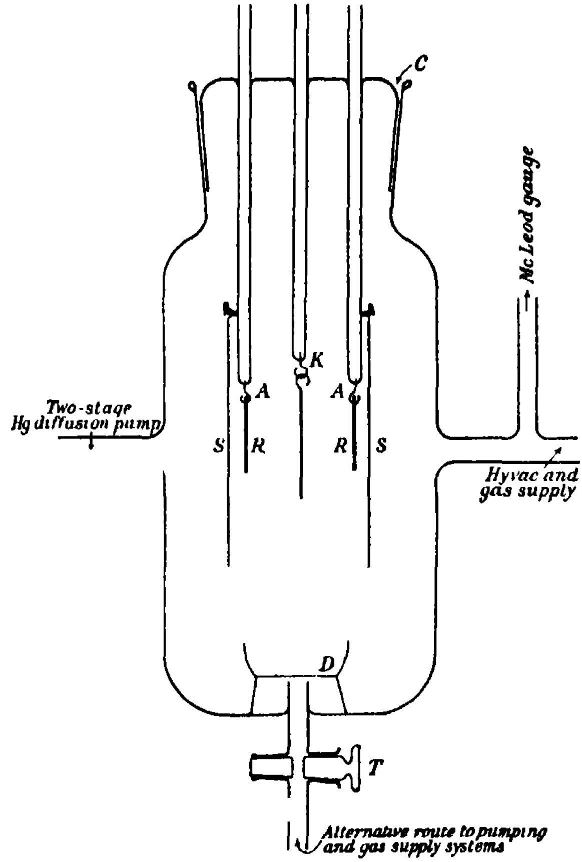

C
McLeod gauge
K
A
A
Two-stage Hg diffusion pump
S
R
R
S
Hyvac and gas supply
D
T
Alternative route to pumping and gas supply systems

Fig. 1.

by running warm low-vapour-pressure grease into the channel, C, where it was allowed to solidify before evacuating the apparatus. In this manner gas evolution due to the action of the discharge upon grease at the joint was successfully avoided. Similarly, the dish, D, supported above the redistilled phosphoric oxide in the bottom of the vessel, protected the grease of the tap, T. Fig. 1 is self-explanatory of the remainder of the sputtering vessel and its supply services.

The reaction vessel, within which the catalytic activity of films sputtered in the above apparatus was measured in terms of the rate of combination at room temperature of electrolytic gas to steam, consisted of a glass vessel somewhat similar to, but smaller than, the sputtering vessel. It was fitted with a wide-bore, closed-limb mercurial manometer and could be connected either to the pumping system or to a supply of previously dried electrolytic gas. The receiver bearing the platinum film under examination was suspended from a hook carried by the stopper of the reaction vessel, the effective volume of which was 464 c.c In the earlier experiments the bottom of this vessel contained redistilled phosphoric oxide. Later, however, for reasons which will be made clear below, this was replaced by distilled water.

The Gases.—Oxygen, nitrogen, and electrolytic gas were prepared and purified in the manner previously outlined.* The purification of argon will be discussed below.

Procedure.—The receivers were cleaned with *aqua regia* and distilled water. With the receivers in position the sputtering vessel and McLeod gauge were repeatedly evacuated and flushed out with the previously dried gas within which sputtering was to be carried out. During the final evacuation the gas pressure was reduced to that chosen for the beginning of the experiment, in the course of which McLeod gauge, sputtering current and discharge voltage readings were taken at frequent intervals. Since the cathode glow was in all experiments abnormal, *i.e.*, the cathode surface was completely covered by the glow, the voltage-current characteristic of the discharge was always positive. Sputtering was carried out under pressures of the order of between 0·04 and 0·5 mm. Discharge current and voltage did not as a rule exceed about 10 mA. and 3000 v. respectively and were usually much lower. The total time of sputtering a film was generally less than 1 hour. On completion of sputtering, either dry air or the gas employed in the experiment was admitted into the vessel to atmospheric pressure, whereupon the receiver with its freshly sputtered film was transferred to the reaction vessel which, after evacuation, was filled as required with dry electrolytic gas, usually to an initial pressure of 30 cm. With a pair of glass or quartz plate receivers, the counterpart was reserved either as a control or for the purposes of electron diffraction examination.

---

* 'Proc. Roy. Soc.,' A, vol. 111, p. 257 (1926); vol. 124, p. 303 (1929).

The Results.

Films Sputtered in Oxygen.—Preliminary experiments showed that platinum films could be sputtered readily in oxygen under a wide range of conditions of pressure and discharge voltage and current. With certain exceptions these films generally proved to be more or less active promoters of the combustion of hydrogen by oxygen at room temperature, but differed in that, whilst some exhibited a pronounced activity immediately on being brought into contact with electrolytic gas, others under similar circumstances were completely inactive until after the lapse of a more or less prolonged period of induction. It was also found that, given a clean cathode and the absence of grease vapours, the chief experimental conditions determining the character of the sputtered films were (i) the gas pressure, (ii) the discharge current or voltage, either of which in conjunction with (i) determined the other, and, finally, (iii) the duration of sputtering Bombardment of grease by the discharge, or the presence of grease vapours resulted in the evolution of gas, and films sputtered under such conditions exhibited an irregular behaviour, owing apparently to poisoning in the catalytic sense. For similar reasons cleanliness of the cathode was found to be essential and could readily be ensured by a suitable running-in, in the manner previously outlined,* prior to an actual sputtering experiment.

In order to define the conditions governing the formation by sputtering in oxygen from a platinum cathode of (i) inductively active, (ii) immediately active, and, finally, (iii) inactive films, it was found necessary to cover a wide range of conditions, which in turn involved the carrying out of many experiments. To illustrate the conditions leading to the production of each of the three types of films, however, in addition to giving a detailed account of a few experiments selected as typical of each case, it will suffice to set forth the general results obtained in the form of brief summaries.

Inductively Active Films.—The conditions of a typical experiment resulting in the formation of an inductively active film were as given in Table I.

The resulting film, the appearance of which was characteristic of thin inductively active films, was faint straw-yellow in reflected daylight and translucent. On being suspended in the reaction vessel in contact with electrolytic gas at a pressure of 30 cm. and at room temperature, it was found to be wholly inactive and continued so without any change in appearance during 42 minutes, whereupon a few minute black spots began to appear and increased steadily in area. During the next 5 minutes the manometer fell to 28·5 cm. thus showing that the film had begun to catalyse the combustion of electrolytic gas simultaneously with the appearance of the black spots. The activity continued to rise until the pressure was falling at a steady rate of 3 mm. per minute. About 2 hours after the first appearance of spots the film had become uniformly sooty black and semi-opaque.

---

* Finch and Cowen, ' Proc. Roy. Soc.,' A, vol. 111, p. 257 (1926).

Table I.

Total duration of sputtering (min.).	Discharge voltage.	Gas pressure (mm.).	Remarks.
0	600	0·33	A low rate of sputtering producing a relatively slight deposit The discharge current was maintained throughout at 1·1 mA.
1	—	0·30	
3	675	0·28	
8	725	0·24	
13	770	0 22	

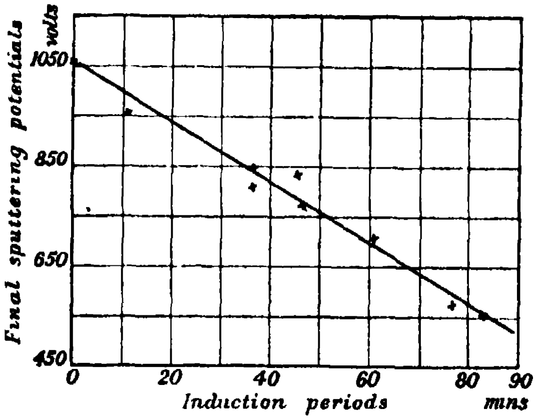


Fig. 2.

Sputtering conditions determining the life of the films in the inductive, *i.e.*, pre-active state, are shown in fig. 2, in which is recorded the relationship between the final discharge voltage and the inductive life of films which were sputtered under approximately similar conditions of final gas pressure. For reasons previously outlined the rate of dissipation of energy at the cathode, and hence also the actual rate of sputtering, decreased with decreasing voltage.* It is clear therefore that, under the conditions of our experiments, a low rate of sputtering in conjunction with a relatively high gas pressure favoured the production of inductive films. On the other hand, it was found that inductive films could also be obtained under conditions involving a high rate of sputtering and therefore entailing relatively high rates of dissipation of energy at the cathode, provided that the total time of sputtering was either short, *i.e.*, of the order of, or less than, 2 minutes, or broken up into short sputtering periods separated by intervals sufficiently long to avoid an excessive rise in the temperature of the sputtering vessel and its contents. Thus, for example, an exceedingly thin film sputtered in 32 seconds with 10·2 mA. at 1200 volts, the final oxygen pressure being 0·38 mm., had an inductive life of 21 minutes, after which it exhibited a low but steady activity corresponding to 0·4 cm. fall of manometer per minute. Another film prepared under closely similar conditions, but with a total time of sputtering of 10 minutes, made up of twenty 30-second sputtering periods separated by intervals of approximately 10 minutes, during each of which the loss in oxygen pressure due to "clean-up" was made good, had an inductive life of 18 minutes and was then exceedingly active (5 cm. per minute). Rapidly sputtered inductive films were found to develop much greater activity than did slowly sputtered films of approximately similar thickness.

The longest inductive life which we have so far been able to observe for films sputtered on to the cylindrical glass receivers was one of 168 minutes. All inductively active films were faint yellow to brownish-yellow in colour, according to the thickness of the deposit, and even the densest of such films were translucent. With the development of activity as a result of a more or less prolonged immersion in electrolytic gas, however, the films invariably darkened and acquired finally either a sooty black or bright metallic mirror surface, this change being accompanied by a remarkable increase in their opacity. It was also found that inductive films blackened and were brought into the active state by contact with moist hydrogen alone. Inductive films were insoluble in *aqua regia* and adhered so firmly to the receiver surface that polishing with a moist abrasive, such as rouge, was necessary for their removal. Once blackened and rendered active, however, whether by contact with electrolytic gas or with hydrogen, the films were not only readily soluble in *aqua regia,* but could also be easily wiped off the receiver with either dry or moistened cotton wool. From these facts the conclusions may be drawn that the inductive deposits consisted in the main of some oxide or, possibly, a mixture of oxides of platinum which were wholly inactive in the sense that they did not, as such, catalyse the union of hydrogen with oxygen, but that reduction by hydrogen to the black and more or less opaque deposit resulted in the formation of an active form of platinum or compound thereof.

---

* Güntherschulze ('Z. Physik,' vol. 36, p. 563 (1926) and vol. 38, p. 575 (1926)) and Von Hippel and Blechschmidt ('Ann. Physik,' vol. 81, p. 1001 (1926)) have shown that the rate of cathodic sputtering is independent of the cathode temperature but is directly proportional to the current, and increases rapidly with either increasing voltage or atomic weight of the gaseous medium within which sputtering is effected

The inductive life of the films could be reduced simply by storage in dry or moist air or *in vacuo* Thus, for example, of two films deposited simultaneously on quartz plates, the one tested immediately on removal from the sputtering vessel was found to have an inductive life of 21 minutes, whereas that of its counterpart after being kept overnight in dry air at atmospheric pressure and room temperature was only 7 minutes

Table II.

Total time of sputtering (no intervals) (secs.).	Sputtering current (mA).	Final voltage.	Final oxygen pressure (mm.).	Cathode temperature at end of experiment	Appearance and behaviour of films in contact with $2H_2 + O_2$ at room temperature.
85	7·0	1175	0·28	Dull red	Thin translucent yellowish-brown inductive film ; blackened and became highly active (3·4 cm. per min.) after 18 mins.
90	8·8	1340	0·21	Red	Thin gold-brown film ; immediately active (2·7 cm. per min.), blackening began immediately in certain areas from which it spread slowly throughout the film.
180	6·9	1010	0·25	Bright red	Slightly opaque, dark brown film , blackened immediately and uniformly over its entire surface and was at once highly active (5 cm. per min.).

Immediately Active Films.—It has been shown above that inductively active films could be prepared in two ways, one of which involved sputtering at a relatively high rate for a short time. Other conditions being equal, a further and sufficient increase in either the rate of cathodic energy dissipation or duration of sputtering resulted in the production of immediately active films. The details, Table II, of three experiments will suffice to illustrate these facts.

Thus, immediately active films could be obtained merely by prolonging the duration of sputtering, the gas pressure and discharge voltage and current being such as would otherwise with a sufficiently short time of sputtering have resulted in the formation of inductively active films. This fact strongly suggests that in all cases the films were actually deposited in the inductive state, and that the production of the immediately active modification was essentially due to the effect of heat emanating from the cathode and acting upon the deposited film. In connection with this it is significant that the life in the latent state of inductive films sputtered under conditions involving relatively low rates of cathodic energy dissipation was independent of the time of sputtering (see fig. 2).

The above view of the manner of the formation of immediately active films was confirmed by the results of experiments of which the following is typical.

A film was sputtered on to a glass tube receiver, the final oxygen pressure, voltage, current and the time of sputtering being 0·24 mm., 770 volts, 1·1 mA. (constant throughout) and 13 minutes respectively. On removal from the discharge vessel, the tube was cut in half. One piece was baked in air at 120° during 30 minutes, whereupon, after cooling, it proved to be immediately active. The other half which had not been heated was then tested and was found to have an inductive life of about 40 minutes.

Immediately active films produced either by sputtering or by the direct heat treatment of inductive films were insoluble in *aqua regia* and adhered firmly to the substrate. Unlike the typically straw to gold-yellow, translucent inductive films, however, immediately active films were darker, brown to grey and less pure in colour and were generally slightly or even sometimes semi-opaque.

Immediately active films formed by sputtering could be kept either *in vacuo* or in dry or moist air during several months and probably longer without complete loss in activity. It was found, however, that when once a film had been completely blackened and its activity thus fully developed by immersion in either electrolytic gas or hydrogen, drying by storage in dry air or by evacuation during half an hour resulted in a total loss of activity, which could not be even partially restored either by heating in the manner previously outlined or by prolonged contact at room temperature with either moist air, hydrogen, electrolytic gas, or water. On the other hand, a blackened film of fully developed activity could be stored in distilled water or in moisture vapour at 12 mm. pressure during a fortnight, and probably longer, without loss in activity.

The continuous working life of a catalytically active film was not indefinitely long, even when precautions were taken to keep the electrolytic gas atmosphere saturated with moisture and free from possible poisons such as traces of hydrogen sulphide or grease vapours.

Inactive Films.—The sputtering conditions found to be most favourable for the production of films which were completely inactive, no matter how long they were left in contact with electrolytic gas at room temperature, were those which resulted in the cathode being maintained at a high temperature. The records in Table III of three experiments will serve to illustrate these facts.

Table III.

Duration of sputtering (mins.).	Current (mA.).	Voltage.	Gas pressure (mm.).	Cathode temperature.
0	4	—	0·32	—
2	4	600	0·25	—
4	4	1300	0·16	Red
8	3·2	1400	0·13	Bright red
9	3·5	1600	0·12	,,
11	2	2000	0·11	White hot

The resulting almost completely opaque film had a deep blue-black, mirror-like surface and was wholly inactive. The second experiment was carried out under similar conditions, but was interrupted after 2 minutes sputtering and yielded a translucent yellow film with an inductive life of about 70 minutes. Finally, in the third experiment, terminated after 8 minutes, an immediately active film was obtained.

These results not only afford further direct support of the view outlined above, according to which platinum sputtered in oxygen forms in the first place an inductively active deposit which is subsequently converted by heat into the immediately active form; but they also suggest that, during sputtering, inactive films are likewise formed from the inductive *via* the immediately active variety, as a result of a more intensive heating. On the other hand, it might be supposed that the inactivity of films, in the case of which the final sputtering pressure approached the limiting gas clean-up pressure, was due to the deposition of platinum unaccompanied by oxygen. The results of experiments, however, of which the following are representative of the conditions employed at once negative any such view.

A film was sputtered in oxygen, the initial and final pressures being 0·35 and the low, nearly limiting "clean-up" pressure 0·10 mm. respectively. In order to reduce thermal effects as far as possible, the current was maintained at the low value of 0·55 mA., the final voltage being 2200 volts. The resulting metallic, blue-grey film was immediately active (1·6 cm. per minute). A second film, prepared under closely similar conditions, was heated during 1 hour at 240° in air and was then found to be completely inactive.

Films Sputtered in Argon.—The argon employed contained between 3 and 4 % of adventitious impurities, 0·2 % being oxygen.

Only immediately active or completely inactive films could be prepared by sputtering in argon. The range of conditions leading to the production of immediately active films corresponded closely to those which were found to yield either the inductive or the immediately active varieties when sputtered in oxygen. Furthermore, the conditions suitable for the formation of inactive deposits were independent of whether sputtering was carried out in argon or in oxygen. It was also found in these experiments that the impurities contained in the argon were removed in the initial stages of sputtering, but that little or no "clean-up" of argon, as such, occurred. These facts may be illustrated by the records of experiments given in Table IV.

Table IV.

Total time of sputtering (mins.).	Current (mA.)	Final voltage.	Gas pressure (mm.).	Remarks
0	2·2	1135	0·43	"Clean-up" of impurities occurred within 10 mins. Cathode remained cool throughout. Sooty black, completely opaque film.
5	2·2	1150	0·42	
10	2·2	1160	0·41	
20	2·2	1155	0·41	
35	2·2	1165	0·41	
60	2·2	1160	0·41	
70	4·4	1350	0·41	
80	4·4	1345	0·41	

The resulting film was immediately and exceedingly active (11·6 cm. per minute). Another film, obtained after 6 minutes sputtering at 1·1 mA. and 3000 volts, the argon pressure being 0·14 mm., had a bright mirror surface and was completely inactive. During this experiment the cathode was at a bright red heat. Finally, in further experiments the relationship between film thickness and maximum activity was determined. For this purpose the deposits were prepared under constant conditions of current (2·2 mA.), (voltage 1170 volts), and argon pressure (0·39 mm.), the time of sputtering, to which the film thickness was directly proportional, alone being varied. The results, incorporated in fig. 3, establish the fact that the activity of these films was directly proportional to their thickness, thus proving that, under the

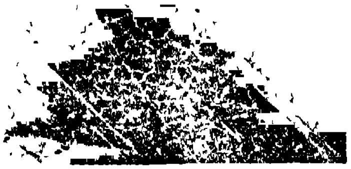


FIG. 6

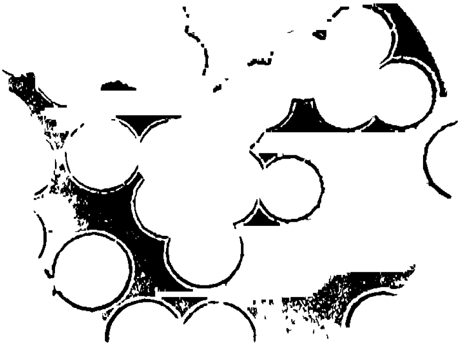


FIG. 5

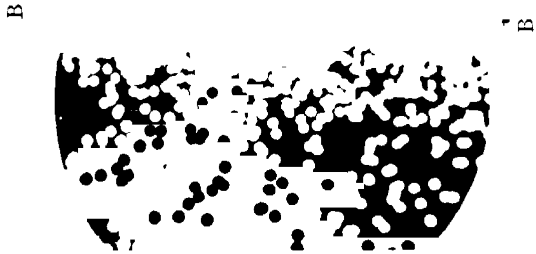


FIG. 4

(*Facing p.* 424.)

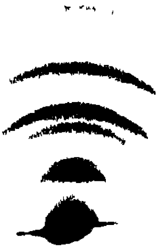


FIG. 7

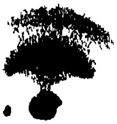


FIG. 9

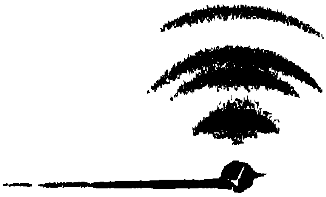


FIG. 8

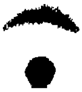


FIG. 10

conditions of our experiments, catalytic action took place throughout the whole film and was not confined merely to the surface. From a colorimetric analysis ($PtI_6^{\ominus\ominus}$)* it was estimated that the average thickness of the film sputtered in 19 minutes (see fig. 3) was of the order of $6 \cdot 4 \times 10^{-6}$ cm., assuming face-centred cubic structure with a cube side of 3·9 A. All films sputtered in argon were readily and completely soluble in *aqua regia.*

Films Sputtered in Nitrogen.—It will suffice to summarize these results as follows.

As with argon, only immediately active and inactive films were obtained. Furthermore, the sputtering conditions suitable for the formation of each type of deposit were similar for either gas. In nitrogen, however, unlike in argon, " clean-up " occurred and continued steadily during sputtering down to a limiting pressure similar to that observed in the case of oxygen; but, other conditions being equal, the rate of such " clean-up " was about four times less than in oxygen. Films sputtered in nitrogen closely resembled in appearance those obtained in argon, forming either opaque, sooty-black deposits or bright, mirror surfaces. They were readily soluble in *aqua regia.*

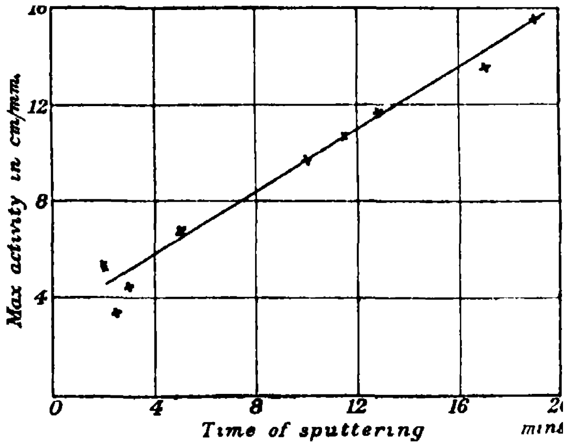


FIG 3.

Microphotographic Examination of Films Sputtered in Oxygen.—The different types of films discussed above all appeared to possess perfectly homogeneous structure when examined under the microscope with a Zeiss 2 mm. apochromatic oil-immersion objective in conjunction with a Swift oil-immersion condenser. Owing to the colour and translucent nature of the inductive and immediately active films obtained in oxygen and the opacity of the active films formed therefrom by contact with hydrogen or electrolytic gas, however, it was possible to observe the microscopic changes occurring during blackening and the development of activity. Whilst under observation on the microscope stage, films sputtered on to glass plates were developed in a hydrogen-air jet, blackening being arrested as desired by removal of the jet. In the case of fig. 4, Plate 6, an immediately active film was deposited on the right-hand half of the slide, whilst the other half was protected by a mica screen. The screen was then removed, and the whole slide, including the active deposit, was covered by sputtering with an inductive film, whereupon development was carried out in the manner outlined above. The immediately active area blackened uniformly within a few seconds, in spite of the relatively thick superposed inductive layer, thus clearly revealing the porous nature of such an inductive film. Meanwhile no change whatever was to be observed in the adjoining purely inductive deposit, except at its boundary with the active film, until towards the end of its inductive life of 21 minutes, as determined by means of the counterpart. Minute circular black spots then appeared and grew steadily, until further development was arrested at the stage shown in fig. 4 (left-hand side) by removing the hydrogen jet. It will be seen that, during its inductive life, blackening was steadily spreading across the boundary, BB, fig. 4, into the inductive deposit, thus proving that catalytic action occurring on active centres can and does bring about the decomposition of that completely inactive platinum oxide forming the main constituent of the inductive type of film to some form of platinum or compound thereof which makes the film catalytically active. We may, therefore, conclude that inductive films owe their eventual activity to the presence of a comparatively few particles of such immediately active material disseminated throughout the deposit. The fact that the inductive film of fig. 4 had a life of 21 minutes in the pre-active state, in spite of the considerable number of active centres revealed on partial development of activity, suggests that the active particles must have been exceedingly small. It has been pointed out above that the pre-active life of an inductive film decreased with the age of the film. It was also found that on storage the number of spots appearing on development likewise increased. In view of these facts it is difficult to resist the conclusion that the platinum oxide slowly decomposes spontaneously to form the immediately active material.

---

* Yoe, " Photo-Chemical Analysis " (Wiley & Sons, New York), vol. 1, p. 355 (1928).

Minute transparent centres are plainly visible in the spots shown in fig. 4. These were apparently caused by contraction and subsequent rupture of the catalytically active film, thus leaving bare the glass substrate in the centre of each spot. With further development, however, the spots gradually healed in the centre, as is shown in the microphotograph, fig. 5, Plate 6, whereupon a characteristic zone of rupture appeared at, and kept pace with the expansion of, the periphery of each active spot. Such peripheral ruptures healed where spots coalesced.

The appearance of fissures, fig. 6, Plate 6, following approximately linear tracks and generally accompanied by peeling of the deposit from off the substrate, signalled the approaching end of the active working life of a film as a catalyst. Such direct evidence of contraction and loss of adhering powers strongly suggest that exhaustion of the activity of platinum films is in some manner closely connected with profound changes in their crystalline structure.

Classification of Sputtered Platinum Films.—The chief types of films obtained in the course of these experiments may now be classified according to their more striking properties as follows :—

Type A.—Inductive films, consisting in the main of a catalytically completely inactive platinum oxide, but containing inclusions of a relatively few particles of an active substance.

Type B.—Differs from Type A only in degree, in that while it still consists largely of the platinum oxide it also contains the active substance disseminated throughout in far greater profusion than in the case of Type A.

Type C.—A catalytically active form or compound of platinum or mixture thereof resulting from the development of either Type A or B in hydrogen or electrolytic gas, and differing from Type B in that its activity is completely destroyed by drying, and thus forms.

Type D.—Differs from Type F in that it shows no signs of fissuring or peeling from the substrate.

Type E.—An inactive form of platinum obtained either by sputtering in oxygen, nitrogen, or argon, or by heating Types A or B.

Type F.—An inactive form of platinum, resulting from the exhaustion of the working life of Type C and characterized by fissuring and loss of adhesiveness to the substrate.

Tests by Electron Diffraction. (G. P. T. and C. A. M.)

In order to discover the crystalline structure of the spluttered films, and to examine any possible correlation between this structure and the catalytic activity, the surfaces were examined by the method of electron diffraction in the apparatus* formerly used by one of us for the study of chemically deposited platinum and a number of other surfaces of various chemical compositions.† The method consists in making use of the wave properties of electrons in the form of cathode rays of about 30,000 volts energy, to derive diffraction patterns from the small crystals of the film, the process being strictly analogous to the Debye-Scherrer method with X-rays, except that, owing to the low penetrating power of cathode rays, the thickness of the film effective in producing the patterns is limited to the order 10^{-6} cm. As in the Debye-Scherrer case, the diffraction patterns take the form of concentric circular rings, or rather semicircles, since about half the field is blocked out by the shadow of the block on which the film has been deposited. From the radii of the rings we can deduce the spacings of the crystal planes which cause them, and see if they agree with those of any known structure. From the width of the rings it is sometimes possible to deduce the size of the small individual pieces of crystal, for if these latter are very small there will be a diminution in their "resolving power" regarded as analogous to that of a grating, and the rings, which are analogous to the spectral lines of the grating, will be widened in consequence. If, however, the dimensions of the crystals exceed about 10^{-6} cm. the resolving power becomes so good that the width due to the crystal size is small compared with that due to the finite size of the electron beam, and to its lack of homogeneity, and nothing can be said as to the size of the crystals except that it is at least of the above order of magnitude.

A surprising variety of patterns was found, excluding a few exceptional cases, they can be classified as follows :—

I. Normal platinum, a pattern characteristic of the normal face-centred cubic platinum crystals with a cube side of 3·91 A. They are usually fairly sharp, but the breadth of the rings has generally been slightly increased by the small size of the crystals, fig. 7, Plate 7.

II. As the above, but with the rings not of uniform intensity, thus showing that the crystals are orientated in preferred directions with respect to the surface of the receiver on which the platinum has been sputtered.‡

III. A pattern consisting of the normal platinum together with some extra rings which have not yet been identified. These latter frequently show orientation even when the platinum rings do not. It is possible that there is more than one group of these " extra " rings. They were always narrow though often faint, fig. 8, Plate 7.

---

* Thomson and Fraser, 'Proc. Roy. Soc.,' A, vol. 128, p. 41 (1930).

† Thomson, *ibid.*, p. 649.

‡ 'Phys. Soc. Proc.,' vol. 45, p. 381 (1933).

IV(*a*). A pattern consisting chiefly of two strong broad rings with a few faint ones sometimes visible outside. There was usually a strong background, fig. 9, Plate 7.

IV(*b*). A pattern characterized by two very strong sharp rings with a number of fainter ones outside them, fig. 10, Plate 7.

Identification of Patterns IV(a) and IV(b).

It was quickly recognized that these two patterns were due to the same substance with a difference in the size of the crystals, which determines the width of the rings. In fact, there is almost a complete gradation between the two, one of the IV(*b*) type having rings as wide as 1·3 mm. and one of the IV(*a*) type as narrow as 1·5 mm. Usually, however, the IV(*b*) type are very sharp (*c.* 0·8 mm.) and IV(*a*) type at least 2·0 mm. wide. The mean values for the spacings producing the two main IV(*a*) type rings are 2·59 and 1·52 A., and those for the two strong rings of the IV(*b*) type are 2·69 and 1·56 A. It is doubtful if any importance is to be attached to the difference. It is possible that the effective spacing for a crystal as small as those giving rise to IV(*a*) (about 6 atom diameters) is less than that of the same matter in bulk, but it is equally likely that it is a physiological effect due to the difficulty of measuring the centre of a diffuse ring on variable background.

As it seemed probable that this pattern was due to the presence of oxygen it was natural to try an oxide of platinum. A specimen of platinum dioxide reputed 99% pure obtained from Messrs. Griffin & Tatlock, was very kindly tested for us in an X-ray diffraction camera by Mr. Sykes, of Metropolitan-Vickers. The result showed a large percentage of metallic platinum in the specimen, which had presumably partially decomposed, but there were also present diffuse bands, due to a substance in a state of minute subdivision, corresponding to spacings 2·65 and 1·56 A. We consider that we are justified in identifying this substance, which is presumably colloidal platinum dioxide, with the substance of similar spacings which causes IV(*a*) and (*b*). It is, of course, impossible to be sure that small quantities of other substances are not present, although they do not appear in the pattern on the plate. This is especially so with a diffuse pattern such as IV(*a*). Thus we do not wish to assert that there is no metallic platinum in the films of type IV(*a*) and (*b*), only that the compound is in excess. Patterns IV(a) and IV(b) have so far only been obtained in oxygen, as might be expected on the above view of their origin.

Table IV.—Spacings of Patterns III and IV(b) in Angstroms.

III	4·4	3·06 s.	2·06 m s.	1·73	1·31	and smaller.				
IV(b)	2·69 s.	1·56 s.	1·35	1·01	0·90	0·775	0·742	0·672	0·610	0·578

also a trace of a diffuse orientated ring with spacing about 2·2.

Patterns I and II.

Pattern I was formed in argon, oxygen, nitrogen, and hydrogen, and is regarded as normal. When formed in oxygen, it is inactive catalytically or only slightly active. Pattern II occurs in argon, oxygen, and nitrogen. In most cases in which it occurred the film had been heated considerably by radiation from the sputtering cathode. In all probability the orientation of the crystals of platinum which causes this pattern is due to their growth in a certain preferred direction being made easy by heat.* There seemed to be no necessary connection between the orientation and the catalytic activity of the films. Some orientated films were active, but only those sputtered in argon in which films of pattern I were commonly active, and the proportion of active films in the two cases was not notably different. In what follows patterns I and II will be treated together.

Films Sputtered in Oxygen.—In this gas it has been found possible to correlate the type of pattern observed with the conditions of the discharge used in preparing the specimen. If the voltage never rose above about 900 volts with a current of 2·2 mA., corresponding to a pressure of about 0·35 mm. or over, type IV(b) resulted, if the voltage was above 900 with the same current, but below 2000, we get type IV(a). The pressure for these was usually between 0·1 and 0·35 mm. Since the *rate* of sputtering is chiefly governed by the voltage, these results are consistent with the view that the large crystals of IV(b) require a slower rate of deposition than the small ones of IV(a). If the voltage was much above 2000 I or II appeared, the cathode being visibly hot and thus heating the receiver by radiation, while III occurs in a narrow range of conditions with the voltage near 2850. With certain exceptions, which will be mentioned shortly, patterns I, II, and III occur on films which are catalytically inactive, while IV(*a*) and (*b*) are always potentially active, in the sense that they can be developed by electrolytic gas or hydrogen into active films. In all 43 films were studied by diffraction: 10 showed I and II; 3 showed III; 17 showed IV(*a*); 7 showed IV(*b*); one showed I on part of its surface and III on the rest, the two parts being distinguished by colour; one was intermediate between IV(*a*) and IV(*b*), one was uncertain, probably IV(*a*); another probably a mixture of I and IV(*a*); in two cases the rings were too poor to enable the pattern to be assigned, the films being very thin.

---

* An account of these orientated films was given before the Phys. Soc., February 3, 1933.

There were two cases in which a film was active which did not show pattern IV.

Two additional facts are of importance. If a film originally of type IV(*a*) or (*b*) is used catalytically till it ceases to be active, and then is re-examined by diffraction, the pattern is found to have completely changed to the normal platinum type I or more rarely to III. If type IV(*a*) is heated to 225° C. *in vacuo*, the same change is produced, and the film is then *inactive*. Films giving patterns IV(*a*) and IV(*b*) are yellow, while I, II, III are usually grey or black, so the change can be observed visually as well as by diffraction.

The counterparts of all films of the type IV(*b*), and of some of those of type IV(*a*), were inductive. The specimens themselves had usually a shorter inductive period than their counterparts and were sometimes active immediately. The reduction or disappearance of the inductive period in the specimens tested by electron diffraction is probably a normal consequence of the time that necessarily elapses before they can be used for catalysis; there is no evidence to suggest that any detectable change occurs in the pattern on keeping. On the contrary, a specimen showed the characteristic IV(*b*) pattern three days after being made.

The mean width of the rings in the IV(*a*) specimens is 2·1 mm. (range 1·5–2·6) which corresponds to a crystal size of about $1{\cdot}6 \times 10^{-7}$. The IV(*b*) specimens had often quite sharp rings and their crystal size must have exceeded 10^{-6} cm. in most cases.

Films Sputtered in Argon.—These showed before catalytic testing only patterns I, II, and III, the latter once only. They were active in 12 cases out of 44 tested.

Films Sputtered in other Gases.—Three films sputtered in nitrogen were inactive. All showed normal platinum structure with slight orientation in two cases. Three others sputtered in hydrogen were also inactive and showed normal platinum structure without orientation.

Pattern III (for spacings see Table IV).

In several cases films prepared in argon showed pattern III after catalytic use; one examined both before and after showed I before, and an abnormal pattern (of "extra" rings alone) after catalytic use. A similar effect was observed in one of the oxygen films (probably IV(*a*) but pattern poor) which showed after partial blackening an abnormal pattern consisting of the "extra" rings only. One of the IV(*a*) type oxygen films changed to III after use, while one of the abnormal oxygen films of pattern I which was slightly active did the same. There is thus some evidence that when a *feebly* active film is used catalytically the "extra" rings of pattern III are formed, though they can certainly also be formed directly by sputtering, while, on the other hand, *strong* activity leaves normal platinum.

Influence of Crystal Size.—It has often been suggested that the activity of platinum is due to a minute stage of subdivision. The particle size, if it is less than about 30 units cells in each direction can be deduced from the increased thickness of the rings due to small crystals. Many of the photographs do, in fact, show this effect* with the rings due to metallic platinum, the most usual size is about 20 unit cells in each direction, but occasionally there are as few as six. Taking the films sputtered in argon, in which gas a number of films showing the pattern of metallic platinum exhibit activity, a comparison shows so far no tendency for the films with smaller crystals to be the more active.

The size of crystals of platinum dioxide does, however, seem at first sight to be correlated with the inductivity. This is probably apparent only, the size of the crystals depending on the rate of sputtering, the inductive period on the temperature of the surface. Usually slow sputtering involves small evolution of energy and low temperature, but by sputtering intermittently one can keep the temperature low and so get inductive films with small crystals IV(*a*).

Acknowledgment.

The high tension generator used with the electron diffraction apparatus was purchased with a grant made by the Royal Society to one of the authors (G. P. Thomson) for which he wishes to proffer his sincere thanks. He also wishes to thank Mr. T. Riches for his assistance with the apparatus.

One of us (Norman Stuart) wishes to thank the Salters Institute for a grant and, later, a Fellowship which enabled him to devote his whole time to this research

---

* 'Proc. Phys. Soc.,' vol 45, p 381 (1933).

Conclusion and Summary (Joint).

We may summarize the results obtained in the following table in which the state of the film as determined from its catalytic properties and method of preparation is correlated as far as possible with the structure found by electron diffraction :—

Type of film.	Structure
A. *Inductively active.* Prepared in O_2 by— (*a*) Rapid intermittent sputtering. (*b*) Slow sputtering.	IV(*a*) Fuzzy PtO_2 rings, considerable background. IV(*b*) Sharp PtO_2 rings, slight background.
B. *Immediately active.* Prepared by— (1) Sputtering in O_2 (2) Sputtering in Ar.	IV(*a*) Fuzzy PtO_2 rings, considerable background; occasionally also I (normal platinum). I. Normal platinum or II orientated platinum.
C. *Actual state of catalysis (blackened film)* Developed from A or B by electrolytic gas or hydrogen.	Could not be examined owing to the destruction of C and its conversion to D by the drying, caused by putting it in a vacuum
D. *Inactive film* obtained by drying C. E. *Inactive film* when obtained by sputtering in O_2, N_2, Ar, or H_2, or by heating " A " or " B." F. *Exhausted* catalytic film	These all showed :— I. Normal platinum. II. Normal platinum orientated. III. Normal platinum with extra rings. Of these II was rare in films sputtered in O_2, and did not occur when types D or F where developed from IV(*a*) IV(*b*).

It is unfortunate that it has not so far been found possible to test the film by diffraction in the act of catalysis. It is just possible that the peculiar pattern of the " extra " rings alone, found on two films tested after they had been blackened for a short time may be of significance. In one case the film was still slightly active after the diffraction experiment. The blackening may not have been complete.

It may be significant that pattern IV(*a*) is usually accompanied by more background scattering than the others. The presence of single atoms or other non-crystalline bodies would produce increased background scattering, but other causes are possible. The background of a film of type D derived from one of type B is always less than that of the parent film, and the size of the crystals is of course greatly increased. We have observed that the PtO_2 films are difficult to move mechanically, unlike the platinum ones; they never show orientation, presumably because the necessary high temperature would dissociate the compound. Films sputtered in argon are easily removed mechanically, they show orientation readily and are easily poisoned. The rarity of mixed patterns, with the exception of III, is striking. Types D, E, and F are distinguished by their different modes of preparation. There is so far no evidence to suggest that they are fundamentally different.

We have refrained from suggesting any mechanism for the act of catalysis. Further experiments are needed to distinguish between various possible explanations, and are being made by one of us (G. I. F.). The work described in the present paper, which has taken two years and a half to complete, covers the main features of the catalytic process as directly observable, and shows the structure of the film at its various stages. We feel justified in expressing the view that the method of electron diffraction is likely to play as important a rôle in the study of the mechanism of heterogeneous catalytic reactions as is played by spectrographic methods in the study of gaseous reactions.

Free Paths and Transport Phenomena in Gases and the Quantum Theory of Collisions. I.—*The Rigid Sphere Model.*

By H. S. W. Massey, Ph.D., Senior 1851 Exhibitioner, Trinity College, Cambridge; and C. B. O. Mohr, B.A., M.Sc., Trinity College, Cambridge, 1851 Exhibitioner, University of Melbourne.

(Communicated by P. A. M. Dirac, F.R.S.—Received March 1, 1933.)

The necessity for the use of quantum mechanics in the theory of atomic phenomena is most clearly manifest in the study of collision processes. Diffraction effects have been observed in the scattering of electrons from crystals* and by atoms,† while the recent developments of molecular ray technique have made it possible to establish the existence of cross-grating spectra in the reflection of molecular beams from crystal surfaces.‡ In view of the importance of wave theory in these phenomena, it is clearly necessary to examine the conditions under which the classical theory of gases must be modified and to determine the nature of the modifications. Such an investigation receives added importance owing to the possibility of experimental test by molecular ray methods. Also, considerable interest is attached to the possibility of direct experimental proof of the Bose-Einstein statistics for neutral atoms and molecules from collision experiments as has already been possible for α-particles.*

---

* G. P. Thomson, "Free Motion in Wave Mechanics."

† J. J. and G. P. Thomson, "Conduction of Electricity through Gases," vol. 2, Chapter III, Camb. Univ. Press (1933).

‡ Fraser, "Molecular Rays," Camb. Univ. Press (1931).

In order to develop the quantum theory of collisions in a form suitable for this purpose, we first discuss the simplest model which bears sufficient resemblance to the actual facts, and so we consider the rigid sphere model for gas atoms. This model has already proved valuable in the classical theory of transport phenomena and has the additional advantage of permitting an exact quantum mechanical solution. It will be seen that the results obtained by the use of this model are of great interest and suggest several new lines of investigation, both experimental and theoretical. Finally, a method for dealing with the general case of any law of force will be discussed.

§ 1. *The Physical Processes Involved and the Classical Formulæ.*

There is a large number of properties of gases which depend on the interaction energies of the atoms, and many of these properties have already been investigated for a variety of gases. We propose to discuss here only those properties which are directly related to collisions between the gas atoms, and of these the coefficients of viscosity and diffusion (including thermal diffusion) have long been known for various gases over a wide range of temperatures. To these we may now add the direct measurement of free paths, of the angular distributions of gas atoms scattered under various conditions, and of the mobilities of positive ions in pure gases. The latter measurements virtually consist in determining the diffusion coefficients of the ions in the gas.

The great value of these investigations is that, when correlated with a satisfactory theory, they lead to a knowledge of the fields of force between gas atoms, information of considerable value for chemistry, as well as providing a check on the theoretical formulæ obtained from quantum mechanical theory of atomic interactions. Laws of force have already been derived by Lennard-Jones for a number of gases with the use of classical theory in connection with observed coefficients of viscosity and diffusion and their variation with temperature.[*] Direct test of these laws of force by means of free path measurements has no significance on classical theory, as the free path depends on the experimental definition of a collision and tends to zero as smaller and smaller deviations are included in the measurements. On the quantum theory this is no longer true, and it is possible to obtain very interesting and valuable information from direct observation of free paths.

---

* Mott, 'Proc. Roy. Soc.,' A, vol. 126, p. 259 (1930); Chadwick, 'Proc. Roy. Soc.,' A, vol. 128, p. 114 (1930); Blackett and Champion, 'Proc. Roy. Soc.,' A, vol. 130, p. 380, (1931).

Before discussing the quantum theory of collisions, we will give the classical formulæ for the various quantities involved, and convert them to a form convenient for quantum mechanical discussion.

The most complete classical theories of transport phenomena are due to Chapman[†] and Enskog,[‡] who obtained the same final results; we shall use Chapman's formulæ throughout. For the coefficient of viscosity η of a simple gas at absolute temperature T, he finds

$$\eta = \frac{5}{4j^3 M^2}\left(\frac{2\pi}{jM}\right)^{3/2} \frac{1+\varepsilon}{\pi R_{11}}, \tag{1}$$

where

$$j = 1/2\kappa T, \tag{2}$$

M is the mass of a gas atom, and κ is Boltzmann's constant. R_{11} is given by

$$R_{11} = \tfrac{1}{2}\int_{-\infty}^{\infty} V^7 Q_\eta\, e^{-\frac{1}{2}jMV^2} dV, \tag{3}$$

where

$$Q_\eta = \pi \int_0^\infty \sin^2\theta \, dp^2, \tag{4}$$

p is the perpendicular distance between the asymptotes of the paths of two atoms while entering on collision and θ is the angle the relative velocity of the atoms is turned through by the collision. V is the relative velocity of the atoms before impact. Q_η has the dimensions of area and may be defined as the collision area effective in viscosity.

The quantity ε depends in a complicated way on collision phenomena in the gas (and hence on the law of interaction), but on the classical theory its value is never greater than 0·017. As quantum modifications are unlikely to alter the magnitude of ε sufficiently to make its calculation important, we shall neglect it in what follows; in a later paper its calculation will be considered using quantum mechanical collision theory.

---

[*] For a summary of this work see chap. X by Lennard-Jones in R. H. Fowler's "Statistical Mechanics," Camb. Univ. Press (1929), and 'Proc. Lond. Phys. Soc.,' vol. 43, p. 461 (1931).

[†] 'Phil. Trans.,' A, vol. 216, p. 279 (1916); vol. 217, p. 115 (1917).

[‡] 'Inaug. Diss.' (Uppsala, 1917).

For the coefficient of diffusion between two gases (distinguished by suffixes 1 and 2) Chapman gives the formulæ

$$D = \tfrac{3}{16}\pi^{1/2}\left(\frac{M_1 + M_2}{jM_1M_2}\right)^{7/2} \frac{1}{(\nu_1 + \nu_2)\,P_{12}}\,\frac{1}{1 - \varepsilon_0}, \tag{5}$$

where M_1, M_2 are the masses of the gas atoms, and ν_1, ν_2 the number of atoms of each per cubic centimetre. P_{12} is given by

$$P_{12} = 2\int_{-\infty}^{\infty} V^5 Q_D(1, 2) \exp.\left\{\frac{jM_1M_2}{M_1 + M_2} V^2\right\} dV, \tag{6}$$

where

$$Q_D(1, 2) = \pi\int_0^{\infty} \sin^2 \tfrac{1}{2}\theta \, dp^2. \tag{7}$$

and is the collision area effective in diffusion.

ε_0, like ε in expression (1), depends in a complicated way on the atomic collisions and also on the relative concentrations of the two gases. It is never very great—on classical theory its maximum value is 0·136—and so the quantum theoretical calculation of ε_0 will be deferred to a later paper.

The coefficient of thermal diffusion of a gaseous mixture cannot be written down in such a simple form, but it involves, in general, the interactions between molecules of the same kind as well as molecules of different kinds. As a consequence of this, the formula includes integrals of the form

$$P_{11} = 2\int_{-\infty}^{\infty} V^5 Q_D(1, 1) \exp.\{-\tfrac{1}{2}jMV^2\}\, dV. \tag{8}$$

The mobilities of positive ions in pure gases are given in terms of the coefficient of diffusion D by means of Langevin's formula*

$$k = eD/\kappa T, \tag{9}$$

where e is the charge on the ion.

Examination of the classical theory shows that the only modifications which are introduced in the general formulæ above arise from the use of the Bose-Einstein, instead of classical, statistics, and this modification can be neglected except at extremely low temperatures and high densities. However, when we apply the formulæ to the consideration of any particular model, the use of the quantum theory of collisions between the gas atoms will affect the values of the cross-sections Q_η, Q_D in the formulæ (4), (7) above. The quantities ε,

---

* 'Ann. Chim. Phys.,' vol. 5, p. 245 (1905).

ε_0 will also be affected, but we neglect this at present for the reasons already stated.

In order to obtain Q_η, Q_D in forms convenient for quantum mechanical treatment, we must change the variable from the impact parameter p to the angle of scattering θ in relative co-ordinates.*

Since we are using relative co-ordinates we may discuss the collisions as if one atom is held at rest. Suppose, then, that we have a stream of N atoms per unit area per second incident with velocity v on this atom. Then the probable number of particles per second crossing a plane perpendicular to the direction of flight with angular momentum between J and J + dJ is

$$2\pi N J dJ/M^2v^2.$$

To obtain the number of particles deflected between angles θ and $\theta + d\theta$, which we write in the form

$$2\pi N I(\theta) \sin\theta \, d\theta,$$

where I(θ) has the dimensions of area, we make use of the fact that J may be expressed as a function of θ. Therefore

$$2\pi N I(\theta) \sin\theta \, d\theta = \frac{2\pi N J}{M^2v^2} \frac{dJ}{d\theta} d\theta.$$

Now J = Mvp, so

$$p\,dp = I(\theta) \sin\theta \, d\theta. \tag{10}$$

Hence

$$Q_\eta = 2\pi \int_0^\pi I(\theta) \sin^3\theta \, d\theta, \tag{11}$$

$$Q_D = 2\pi \int_0^\pi I(\theta) \sin^2 \tfrac{1}{2}\theta \sin\theta \, d\theta, \tag{12}$$

while the collision cross-section Q is given by

$$Q = 2\pi \int_0^\pi I(\theta) \sin\theta \, d\theta. \tag{13}$$

§ 2. *The Quantum Theory of Collisions.*

The function I (θ) for the scattering of a particle of mass m and velocity v by a field of force of potential V (r) is given by

$$I(\theta) = \frac{1}{4k^2} \left|\sum_n (2n+1)(e^{2i\delta_n} - 1) P_n(\cos\theta)\right|^2, \tag{14}$$

---

* θ is the angle through which the direction of the relative velocity is turned.

where $k = 2\pi mv/h$ and the phases δ_n are obtained from the asymptotic form of the solution of the equation

$$\frac{d^2u}{dr^2} + \left\{k^2 - \frac{8\pi^2 m}{h^2} V(r) - \frac{n(n+1)}{r^2}\right\} u = 0, \tag{15}$$

which is finite at the origin. The phases are such that this asymptotic form is*

$$u \sim \sin(kr - \tfrac{1}{2}n\pi + \delta_n). \tag{16}$$

The collision cross-section will then be given by

$$Q = \frac{4\pi}{k^2} \sum_n (2n+1)\sin^2 \delta_n. \tag{17}$$

Before proceeding to discuss those properties of the phases which are important for our purpose, it is necessary to remark the modifications of treatment necessary when the colliding systems are of comparable mass. In this case we obtain the same expression as (14) in the co-ordinate system in which the position of one atom is defined relative to the other atom, but we must take for the mass m the reduced mass

$$\frac{M_1 M_2}{M_1 + M_2},$$

where M_1, M_2 are the masses of the colliding atoms. $V(r)$ is now, of course, the interaction energy of the two atoms and v their relative velocity. In experimental observations of angular distributions, one measures the number of atoms scattered in a given direction relative to the direction of incidence. To convert the angular distribution in relative co-ordinates to the angular distribution relative to the direction of incidence, one merely uses the classical momentum and energy relations. In particular, if the atoms are of equal mass, the distribution per unit angle will be given by

$$I(2\Theta)\sin 2\Theta, \tag{18}$$

where Θ is the angle of scattering referred to the direction of incidence.

The most important property of the phases is that δ_n is small when

$$\frac{8\pi^2 m}{h^2} V(r) \ll \frac{n(n+1)}{r^2}, \tag{19}$$

for such r that $kr \sim n + \tfrac{1}{2}$.

---

* Faxen and Holtsmark, 'Z. Physik,' vol. 45, p. 307 (1927).

Under these conditions it is possible to use an approximate expression for δ_n in the form*

$$\delta_n = \frac{4\pi^3 m}{h^2} \int_0^\infty \mathrm{V}(r) \{\mathrm{J}_{n+\frac{1}{2}}(kr)\}^2 r \, dr. \tag{20}$$

In the special case where all the phases are small, the series (14) may be summed to give, approximately,

$$\mathrm{I}(\theta) = \frac{64\pi^4 m^2}{h^4} \left\{ \int_0^\infty \mathrm{V}(r) \frac{\sin(2kr \sin\frac{1}{2}\theta)}{2kr \sin\frac{1}{2}\theta} r \, dr \right\}^2, \tag{21}$$

which is the well-known approximation due to Born.† Substituting this formula in (13) we see that the total collision cross-section Q will be finite if $\mathrm{V}(r)$ vanishes at infinity faster than r^{-3}. This same result must hold for the exact formula (14), for, when n is sufficiently large, the exact and approximate series converge together by virtue of (19) and (20). As it is extremely unlikely that the interaction between atoms falls off as slowly as r^{-3} for large r, we see that the mean free path has a perfectly definite value depending on the law of force and so provides a further means of determining this law.

In the particular case of the low velocity limit of the cross-section, it is not yet possible to state the conditions under which the limit is finite, but it appears that, for a potential which vanishes more rapidly than r^{-3} at infinity, the limit is infinite only when very special relations are satisfied by the field of interaction. This is illustrated in the appendix for the case of an exponential field of force.

In the case of the collisions of gas atoms, it is easy to see that Born's approximation is only applicable to phases of very high order except for collisions at extremely low temperatures. As the ratio of wave-length to atomic diameter is considerably less than unity, an approximation based on classical theory will give satisfactory results for the phases of low order. Such an approximation is that given by Jeffreys.‡ This method gives for the solutions of the equation

$$\frac{d^2u}{dr^2} + \left\{ k^2 - \frac{8\pi^2 m}{h^2} \mathrm{V}(r) - \frac{n(n+1)}{r^2} \right\} u = 0,$$

the asymptotic forms

$$\begin{aligned} u \sim \sin\left[\frac{\pi}{4} + \int_{r_0}^\infty \{f(r)\}^{\frac{1}{2}} dr\right], \\ \sin\left[\frac{\pi}{12} + \int_{r_0}^\infty \{f(r)\}^{\frac{1}{2}} dr\right]. \end{aligned} \tag{22}$$

---

* Mott, 'Proc. Camb. Phil. Soc.,' vol. 25, p. 304 (1929).

† 'Z. Physik,' vol. 38, p. 803 (1926).

‡ 'Proc. Lond. Math. Soc.,' vol. 23, p. 428 (1924).

Here

$$f(r) = k^2 - \frac{8\pi^2 m}{h^2} \mathrm{V} - \frac{n(n+1)}{r^2} \tag{23}$$

and r_0 is the largest zero of $f(r)$.

The first of these two solutions is zero at the origin and is the solution we require. Comparing (22) with (16) we obtain

$$\delta_n = \tfrac{1}{2}n\pi + \tfrac{1}{4}\pi + \int_{r_0}^{\infty} [\{f(r)\}^{\frac{1}{2}} - k]\, dr. \tag{24}$$

This approximation is satisfactory when

$$\frac{8\pi^2 m}{h^2} \mathrm{V}(r)$$

is large compared with the centrifugal force term $n(n+1)/r^2$, and is strictly accurate in the classical limit of $h \to 0$. If, then, we wish to calculate the scattering of one gas atom by another, we may calculate the phases for small n by Jeffreys' method and for large n by Born's method. The former method will be accurate when δ_n is greater than unity, the latter when δ_n is less than unity. The intermediate phases may then be obtained by interpolation.

By using this method it is thus possible to consider various types of interactions between atoms.* However, before proceeding to such a detailed investigation, we will consider the collision of rigid spheres, for which we may readily find exact expressions for all the phases.

§ 3. *The Effect of Symmetry.*

Before expressing the formulæ (4), (7) in terms of the phases δ_n we must introduce a modification of the above formulæ which is necessary when the colliding atoms are similar. In this case it is impossible to distinguish experimentally between the incident and struck atoms, and the wave function describing the motion must satisfy certain symmetry properties with respect to the co-ordinates of the two atoms. In particular, if the atoms obey the Bose-Einstein statistics, the wave function describing their motion must be symmetric in the co-ordinates of the two atoms.

Since interchange of the atoms changes θ into $\pi - \theta$ we must take the scattered amplitude in this case in the form

$$2^{-\frac{1}{2}}\{f(\theta) + f(\pi - \theta)\},$$

---

* A preliminary account of this work was given in 'Nature,' vol. 130, p. 276 (1932). The cross-sections given there should be doubled.

where

$$f(\theta) = \frac{1}{2ik} \sum_n (e^{2i\delta_n} - 1)(2n + 1) P_n(\cos\theta).$$

This gives, for the function I(θ), the form

$$I(\theta) = \frac{1}{2k^2} \left|\sum_n (4n + 1)(e^{2i\delta_{2n}} - 1) P_{2n}(\cos\theta)\right|^2,$$

all odd harmonics being excluded. This modification has the effect of reducing the number of effective terms in the series and making the deviations from classical theory more marked at a particular energy of the particles than they otherwise would be.

We are now in a position to express the formulæ (4), (7) in terms of the phases δ_n. Firstly for the case of the viscosity, by using the formula

$$x^2 P_n(x) = \frac{(n+2)(n+1)}{(2n+3)(2n+1)} P_{n+2}(x) + \frac{4n^3 + 6n^2 - 1}{(2n+1)(2n-1)(2n+3)} P_n(x) + \frac{n(n-1)}{(2n+1)(2n-1)} P_{n-2}(x),$$

we obtain

$$Q_\eta = \frac{4\pi}{k^2} \sum_n \left\{ \frac{4n^3 + 6n^2 - 2n - 2}{(2n-1)(2n+3)} \sin^2\delta_n - \frac{2(n+2)(n+1)}{2n+3} \cos(\delta_n - \delta_{n+2}) \sin\delta_n \sin\delta_{n+2} \right\}. \quad (25)$$

Then for the diffusion, using the formula

$$x P_n(x) = \frac{n+1}{2n+1} P_{n+1}(x) + \frac{n}{2n+1} P_{n-1}(x),$$

we have

$$Q_D(1, 2) = \frac{2\pi}{k^2} \Sigma \{(2n + 1) \sin^2\delta_n - 2(n + 1) \cos(\delta_n - \delta_{n+1}) \sin\delta_n \sin\delta_{n+1}\}. \quad (26)$$

In the special case of self-diffusion, which is important in the theory of thermal diffusion, the odd phases vanish and

$$Q_D(1, 1) = \frac{4\pi}{k^2} \sum_n (4n + 1) \sin^2\delta_{2n},$$
$$= \tfrac{1}{2} Q. \quad (27)$$

These formulæ are to be compared with the formula (17) for the total cross-section Q.

§ 4. *The Elastic Sphere Model.*

If we consider the atoms as elastic spheres, we take the interaction energy $V(r)$ as given by

$$\begin{aligned} V(r) &= 0, \quad r > r_0 \\ &= \infty, \quad r < r_0. \end{aligned} \tag{28}$$

Under these conditions the wave function representing the relative motion must vanish at the boundary $r = r_0$. The wave equation from which to determine the phases δ_n is

$$\frac{d^2u}{dr^2} + \left\{ k^2 - \frac{n(n+1)}{r^2} \right\} u = 0, \tag{29}$$

for $r > r_0$. The solutions of this equation are, in terms of Bessel functions,

$$u = r^{\frac{1}{2}} J_{n+\frac{1}{2}}(kr), \quad r^{\frac{1}{2}} J_{-n-\frac{1}{2}}(kr),$$

so that the general solution is

$$r^{-\frac{1}{2}}u = A\, J_{n+\frac{1}{2}}(kr) + B\, J_{-n-\frac{1}{2}}(kr). \tag{30}$$

In order that this solution be zero at $r = r_0$ we must have

$$\frac{B}{A} = -\frac{J_{n+\frac{1}{2}}(kr_0)}{J_{-n-\frac{1}{2}}(kr_0)}. \tag{31}$$

Since

$$J_{n+\frac{1}{2}}(kr) \sim r^{-\frac{1}{2}} \sin(kr - \tfrac{1}{2}n\pi),$$

$$J_{-n-\frac{1}{2}}(kr) \sim r^{-\frac{1}{2}} \cos(kr - \tfrac{1}{2}n\pi),$$

we thus see that

$$\delta_n = \arctan(-1)^{n+1} \frac{B}{A}, \tag{32}$$

and in particular

$$\delta_0 = -kr_0. \tag{33}$$

In Table I the phases δ_n are tabulated for various values of kr_0. The number of multiples of π to add to the smallest solution of (32) for δ_n is determined from the number of zeros of the function $J_{n+\frac{1}{2}}(kr)$ eliminated by the field. As the existing tables of half order Bessel functions* were inadequate for our purpose, a number had to be calculated using the recurrence formulæ for the Bessel functions; it would be sufficient in many cases to use the asymptotic expressions for Bessel functions of large argument† in the relation (31).

---

* Watson, "The Theory of Bessel Functions," Camb. Univ. Press (1922). In using these tables it must be remembered that the definition of $J_{-n-\frac{1}{2}}(kr)$ given by Watson is $(-1)^n$ that given above.

† Watson, chaps. VII and VIII.

Table I.—Values of the phases $-\delta_n$ which occur in the quantum theory of the interaction of hard spheres for different values of kr_0 where r_0 is the sum of the radii of the spheres and $kh/2\pi M$ their velocity, M being the "reduced mass."

n	$kr_0 =$					
	30	20	10	5	3	2
0	30 00	20 00	10 00	5 00	3 00	2 00
1	28 46	18 48	8 53	3 63	1 75	0 89
2	26 96	17 01	7 16	2 47	0 84	0 26
3	25 49	15·59	5 89	1 50	0 28	0 04
4	24 05	14 22	4 73	0 79	0 06	0 003
5	22 65	12 90	3 68	0 32	0 007	
6	21 28	11 63	2 74	0 09		
7	19 94	10 42	1 93	0 02		
8	18 64	9 26	1 25	0 002		
9	17 37	8·15	0 72			
10	16 14	7 11	0 36			
11	14 95	6 11	0 14			
12	13 79	5 19	0 04			
13	12 66	4 32	0·01			
14	11 57	3 52				
15	10 53	2 79				
16	9 52	2 14				
17	8 55	1 56				
18	7 63	1 08				
19	6 74	0 68				
20	5 90	0 39				
21	5 10	0 21				
22	4 34	0 09				
23	3 65	0 03				
24	2 99	0 01				
25	2 37					
26	1 86					
27	1 39					
28	0 98					
29	0 66					
30	0 40					
31	0 24					
32	0 12					
33	0 05					
34	0 02					

The quantum theory formulæ for the hard sphere are then obtained by substituting the expression (32) for the phases in (17), (25), (26). The corresponding classical formulæ are

$$Q = \pi r_0^2, \tag{34}$$

$$Q_\eta = \frac{2\pi}{3} r_0^2, \tag{35}$$

$$Q_D = \tfrac{1}{2}\pi r_0^2. \tag{36}$$

In fig. 1 and Table II the classical and quantum theoretical values for Q_η, Q_D and the collision radius for the case of a hard sphere are compared. In all cases we notice that when the colliding particles are similar, there are deviations of 7% from classical theory for the first of these expressions when the wave-length is greater than one-fifth of the diameter of the atoms, but when the colliding atoms are unlike, the classical theory holds to this degree of accuracy up to wave-lengths as great as one-third of the sum of the atomic radii. Referring to the corresponding temperatures indicated in the figure for the collision of helium atoms, we see that, for light atoms, the

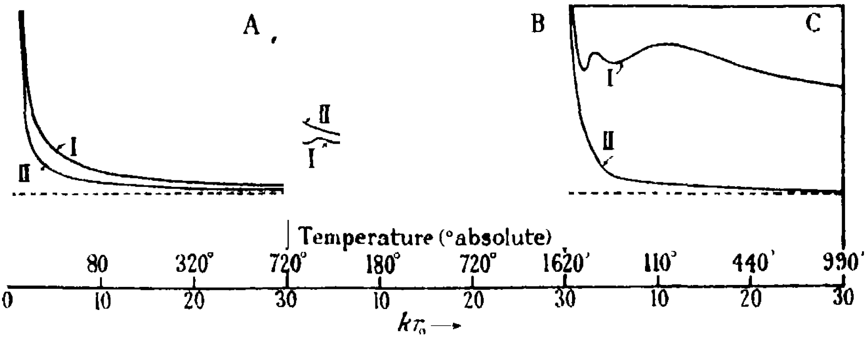


Fig. 1.—Illustrating the ratio of the quantum theoretical to the classical effective collision areas for the viscosity, collision radius, and diffusion, at different temperatures using the hard sphere model. The temperatures refer to helium atoms of diameter 2·1 A. I, Identical atoms ; II, dissimilar atoms ; – – – – – classical value A, collision effective in viscosity ; B, collision radius effective in scattering , C, collision effective in diffusion.

Table II.—Values of the ratio of the quantum theoretical to the classical areas effective in the viscosity, the collision radius, and the diffusion for the rigid sphere model, for different values of $kr_0 = 2\pi r_0/\lambda$ (λ the wave-length). I for identical atoms , II for different atoms.

kr_0	0	2	3	5	10	20	30	∞
Viscosity—								
I	16·00	2·28	1·76	1·47	1·23	1·11	1·07	1·00
II	8·00	2·17	1·78	1·41	1·14	1·06	1·04	1·00
Collision radius—								
I	2·83	1·52	1·58	1·55	1·61	1·53	1·48	1·41
II	2·00	1·73	1·66	1·62	1·55	1·50	1·46	1·41
Diffusion—								
I	8·00	2·30	2·50	2·40	2·59	2·34	2·19	2·00
II	4·00	1·65	1·42	1·17	1·11	1·04	1·02	1·00

deviations from the classical theory are quite important from ordinary temperatures (300° K.) downwards.

For cross-sections we see that the calculated cross-section never tends to the classical, but to a value roughly twice this. It may be proved that this factor of 2 is to be expected without any calculation of the phases as shown in § 5.

§ 5. *Free Paths and Angular Distribution of Scattered Gas Atoms.*

It was seen in § 4 that the quantum theoretical collision cross-section for elastic spheres tends to twice the classical cross-section as the wave-length decreases (*i.e.*, increasing temperature). This may be proved as follows.

When the wave-length is sufficiently small a large number of terms of the series (17) is required and nearly all the important phases are large. We have

$$Q = \frac{4\pi}{k^2} \sum_n (2n+1) \sin^2 \delta_n,$$

and under the conditions stated we may replace the sum by an integral to give

$$Q = \frac{8\pi}{k^2} \int_0^X x \sin^2 \{f(x)\}\, dx.$$

Owing to the magnitude of $f(x)$ we may replace $\sin^2 \{f(x)\}$ by its mean value of $\frac{1}{2}$ giving

$$Q = \frac{2\pi X^2}{k^2}.$$

If the field falls off very sharply at a point $r = r_0$ we have from (33)

$$X = kr_0,$$

so that

$$Q = 2\pi r_0^2. \tag{37}$$

Actually the cross-section may be slightly greater than this owing to the assumption of a definite limit for X. It is clear from this result that the quantum collision area for even, say, billiard balls is still twice the classical, but the difference between classical and quantum theory is confined to such small angles of deviation that the difference is of no practical value in such cases. It is just as if the rigid spheres can, owing to their wave nature, affect each other without interacting in the ordinary sense. With gas kinetic collisions, the angles at which the deviations from classical theory are important, are sufficiently large (of the order of several degrees for light atoms) to be within the reach of experiment, and the evidence in this direction will be discussed below. These conclusions will not be affected by the existence of a weak attractive field, when $r = r_0$ will refer to the distance at which the repulsive field becomes important, Q having the perfectly definite value above. If the attractive field is strong enough to introduce additional phases in the series (17), the cross-section Q may be considerably greater than $2\pi r_0^2$ if r_0 is taken as the distance at which the repulsive field begins to predominate.

From the point of view of the experimentalist it is important to be able to determine the angular resolution necessary in any experiment in order to measure free paths accurately. In particular we need some knowledge of the distribution at small angles where quantum and classical theory differ so markedly. We have

$$\mathrm{I}(\theta) = \frac{1}{4k^2}\left|\sum_n (e^{2i\delta_n} - 1)(2n+1)\,\mathrm{P}_n(\cos\theta)\right|^2, \tag{38}$$

$$= \frac{1}{4k^2}\left\{\left|\sum_n 2\sin^2\delta_n (2n+1)\,\mathrm{P}_n(\cos\theta)\right|^2 + \left|\Sigma_n \sin 2\delta_n (2n+1)\,\mathrm{P}_n(\cos\theta)\right|^2\right\} \tag{39}$$

In the limit of small angles

$$\mathrm{P}_n(\cos\theta) \to 1,$$

and under the conditions of experiment a large number of terms of the series in expression (39) are required, and the δ_n oscillate rapidly with n. Hence

$$\sum_n 2\sin^2\delta_n (2n+1) \gg \sum_n \sin 2\delta_n (2n+1),$$

and so

$$\mathrm{I}(0) \simeq k^{-2}\left|\Sigma (2n+1)\sin^2\delta_n\right|^2,$$
$$= \frac{k^2\mathrm{Q}^2}{16\pi^2} \tag{40}$$

or, in the special case of rigid spheres of radius r_0,

$$\mathrm{Q} \simeq 2\pi r_0^2,$$

so that

$$\mathrm{I}(0) \simeq \tfrac{1}{4}k^2 r_0^4, \tag{41}$$

giving the value at the origin.

At large angles the scattering is classical and we have for a sphere of radius r_0,

$$\mathrm{I}(\theta) = \tfrac{1}{4}r_0^2. \tag{42}$$

This classical formula fails at angles less than the first zero of $\mathrm{P}_n(\cos\theta)$ (which is nearly at π/n) where n is the harmonic of highest order effective in the scattering. n is approximately equal to kr_0. To obtain a sufficiently accurate form for $I(\theta)$ it is only necessary to assume a linear variation at small angles between the value $\frac{1}{4}k^2r_0^4$ at $\theta = 0$ and the classical value which is taken to fail at an angle θ equal to π/kr_0. By multiplying this angular distribution by $\sin\theta$, the error made in counting only deviations greater than a certain angle as indicating collisions can be easily estimated. Fig. 2 illustrates the procedure to be adopted.

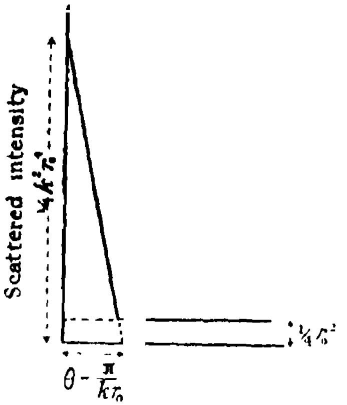


FIG 2.—Illustrating a simple approximate method of obtaining the form of the angular distribution of the scattering of hard spheres of radius r_0 and velocity $kh/2\pi M$, M being the "reduced mass."

In fig. 3 two angular distributions are illustrated for the collision of hard spheres when the wave-length is approximately one-third of the diameter. Curve A is for dissimilar atoms, and curve B for the collision of similar atoms Comparison with the classical curve reveals the behaviour discussed above. Direct experimental test of the form of these curves would be difficult, as any inhomogeneity in the colliding atomic beams would certainly obscure the maxima and minima of curve A. For collisions between similar atoms, there is more hope of experimental verification, as the maximum at 90° (45° when measured relative to the direction of incidence) does not vary in position with relative velocity of the gas atoms. It should therefore be possible to detect it experimentally by scattering one atomic beam by another. Such measurements offer a means of direct proof of the applicability of the Bose-Einstein statistics, for the maximum at 90° is a consequence of these statistics

§ 6. *The Viscosity of Helium and Hydrogen with the Rigid Sphere Model.*

Owing to the independence of temperature exhibited by the classical expression for Q_η, the rigid sphere model predicts on classical theory a variation of viscosity with temperature of the form

$$\eta \propto T^{-\frac{1}{2}}, \tag{43}$$

whereas the experimental evidence* indicates a variation of the form

$$\eta \propto T^{-0\cdot647} \text{ for helium,} \tag{44}$$

$$\propto T^{-0\cdot695} \text{ for hydrogen.} \tag{45}$$

---

* 'International Critical Tables,' vol. 5, p. 2 (1929).

The difference of these exponents from 0·5 was then explained by Lennard-Jones* as caused by deviations from the rigid sphere model, and from them he determined the repulsive field of force between the molecules concerned. If, however, we refer to fig. 1, we see that the quantum theoretical formula gives a

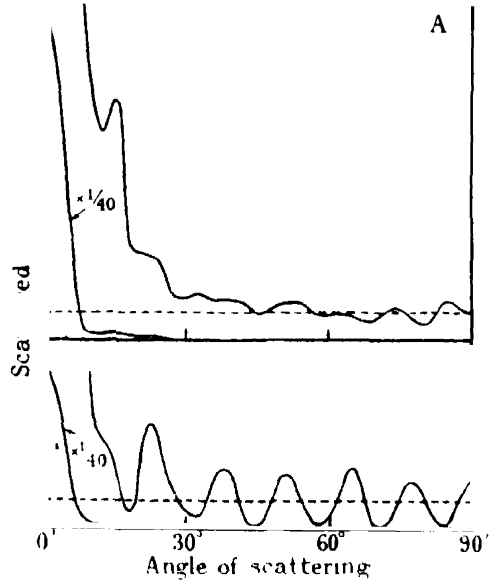


FIG. 3.—Angular distributions (in relative co-ordinates) of helium atoms scattered in helium for the case of $kr_0 = 2\pi r_0/\lambda = 20$. - - - denotes the classical value A, dissimilar atoms; B, identical atoms

variation of η with temperature more rapid than that given by a $T^{-\frac{1}{2}}$ law, and in Table III and fig. 4 the viscosity of helium and hydrogen is shown as calculated on the assumption that the molecules are rigid spheres of diameter

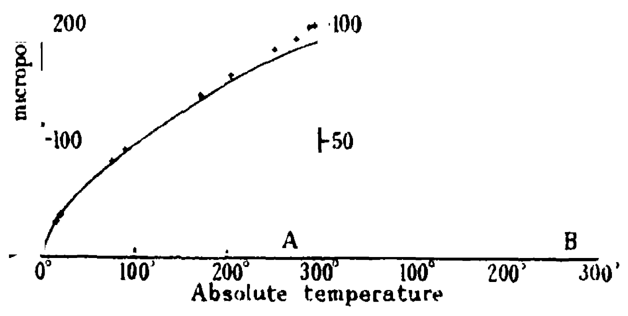


FIG. 4.—Comparison of calculated and observed viscosity of helium and hydrogen at different temperatures. ——— calculated curve, × × × × observed values. A, helium; B, hydrogen.

---

* *Vide* Fowler. *loc. cit.*

Table III.—Comparison of the experimental values of the viscosity of helium and hydrogen at different temperatures with the values calculated on quantum and classical theory using the hard sphere model. (Values of the viscosity are in micropoise.)

	Absolute temperature	Experimental value.	Quantum theory.	Classical theory
Helium	294·5	199·4	185	200
	273·1	187·0	177	193
	250·3	178·8	169	184
	203·1	156·4	150	167
	170·5	139·2	135	152
	88·8	91·8	92	110
	75·1	81·5	81·5	101
	20·2	35·03	35·5	52
	15·0	29·46	30	45
Hydrogen	296·1	88·2	78	88
	273·1	84·2	74	84
	170·2	60·9	56	66
	89·6	39·2	38	48
	70·9	31·9	33	43
	20·6	8·5	15	23
	15·4	5·7	12	20

2·10 A. and 2·75 A. respectively. It will be seen that this model fits the observations for helium over the whole range of temperature within 7%, as contrasted with the classical rigid sphere model which is in error by 50% over this range. This result is very surprising, and seems to indicate that the rigid sphere model is very near the truth, or that viscosity phenomena are not sensitive to the actual fields of force between molecules in collisions when they are properly treated on a wave mechanical theory—at least so far as light atoms are concerned. For heavy atoms (except at very low temperatures) the classical and wave theories tend to the same result, as is evident from fig. 1.

For hydrogen, the agreement is not so satisfactory at low temperatures, but as we are dealing here with a much more complicated phenomenon, the interaction of diatomic molecules, this is not surprising.

In order to test these results further, it is possible to compare the diameters of the molecules obtained from the viscosity with those obtained from the free paths measured by Knauer.* To do this we must use the value of the quantum cross-section in formula (13) for the free path. We then obtain the following comparison between theory and experiment.

---

* 'Z. Physik,' vol. 80, p. 80 (1932).

Table IV.

	Diameter of molecule from viscosity	Effective diameter in collisions	
		Calculated (quantum theory)	Observed (from free paths).
Helium	2 10 A	2 97 A. (2 10 × $\sqrt{2}$)	2 74 A
Hydrogen	2 75 A	3 80 A. (2 75 × $\sqrt{2}$)	3 54 A

The observed values are obtained from Knauer's free path l by using the expression

$$l = \frac{1}{\sqrt{2}\nu Q}, \tag{46}$$

where ν is the number of atoms per cm.^{3} at the pressure of the gas in which the measurements are made. The agreement is good and seems to provide further support for the rigid sphere model with the dimensions indicated. However, owing to the complexity of the conditions of Knauer's experiment, it will be necessary to obtain further measurements of free paths under definite conditions before any decision can be arrived at.

§ 7 *Diffusion and Thermal Diffusion.*

Referring to fig. 1, we expect little deviation from classical theories of diffusion except at very low temperatures. For self-diffusion the identity of the colliding atoms would reduce the classical result by a factor of nearly 2. Although this phenomenon is not observable, the same effect will appear in the thermal diffusion in which integrals of the type of Q_D (11) occur as well as integrals of the type Q_D (12), and this must therefore be taken into account in all calculations of fields of force deduced from observations of thermal diffusion effects.

For the mobilities of positive ions in pure gases, we expect, then, no deviations from classical theory unless the ions are passing through a gas containing neutral atoms with the same nuclei as the ions, as, for example, in the mobility of helium ions in helium. Owing to "umladung"* an appreciable percentage of the ions may be scattered towards large angles and the phenomenon then becomes comparable with self-diffusion. One would thus expect appreciable deviations from classical theory. The calculations for this case are in progress.

* *Vide*, Kallman and Rosen, 'Z Physik,' vol. 64, p. 808 (1930).

It is clear, however, that the most interesting field for investigation, both experimental and theoretical, lies in the comparison of observed and calculated mean free paths, and it is to be hoped that accurate measurements will soon be obtainable for various gases.

Appendix.

Calculation of the Zero Velocity Limit of the Collision Cross-Section for an Exponential Field of Force.

We take the mutual potential energy of the two particles as

$$V = De^{-2ar},$$

The equation for the zero order scattered wave, which alone differs from a Bessel function at the low velocity limit, is in relative co-ordinates

$$\frac{d^2u}{dr^2} + \frac{4\pi^2M}{h^2}(E - De^{-2ar})\,u = 0, \tag{47}$$

where u must satisfy the boundary conditions

$$\begin{aligned} u &= 0 \text{ at } r = 0, \\ u &\sim \sin(kr + \delta), \quad r \to \infty. \end{aligned} \tag{48}$$

Using the substitution $y = e^{-ar}$ reduces the equation (47) to

$$\frac{d^2u}{dy^2} + \frac{1}{y}\frac{du}{dy} + \left(\frac{4\pi^2M}{a^2h^2}\frac{E}{y^2} - \frac{4\pi^2MD}{a^2h^2}\right) = 0. \tag{49}$$

The solutions of this equation are the Bessel functions

$$J_{\pm ik}(imy),$$

where

$$k = \sqrt{\frac{\pi^2ME}{a^2h^2}}, \quad m = \sqrt{\frac{4\pi^2MD}{a^2h^2}}$$

In order to satisfy the boundary conditions at the origin we must have

$$u = AJ_{ik}(imy) + BJ_{-ik}(imy),$$

where

$$AJ_{ik}(im) + BJ_{-ik}(im) = 0,$$

giving

$$\frac{A}{B} = -\frac{J_{-ik}(im)}{J_{ik}(im)}. \tag{50}$$

We have further to break this solution into an incident and a scattered wave.

The asymptotic expansion of u for large r will be obtained from the series expansion of the Bessel functions, which gives

$$u \sim \frac{\mathrm{A}\,(\tfrac{1}{2}ime^{-ar})^{ik}}{\Gamma\,(1+ik)} + \frac{\mathrm{B}\,(-\tfrac{1}{2}ime^{-ar})^{-ik}}{\Gamma\,(1-ik)}. \tag{51}$$

This must be equivalent to

$$\mathrm{C}\left\{\frac{\sin kr}{k} + \alpha e^{ikr}\right\}, \tag{52}$$

where $4\pi|\alpha|^2$ gives the scattering cross-section. Equating the expressions (51) and (52) we find

$$4\pi\,|\alpha|^2 = \frac{\pi}{a^2k^2}\left|\frac{(\tfrac{1}{2}im)^{-ik}\Gamma\,(1+ik)\,\mathrm{J}_{ik}\,(im)}{(\tfrac{1}{2}im)^{ik}\Gamma\,(1-ik)\,\mathrm{J}_{-ik}\,(im)} - 1\right|^2. \tag{53}$$

For a repulsive field m is real and the above expression tends to a finite limit under all conditions. When m is large (as with gas atoms) we obtain, using the asymptotic expressions for the Bessel functions,

$$4\pi\,|\alpha|^2 = \frac{4\pi}{a^2}\,(\log\tfrac{1}{2}m + \gamma)^2, \tag{54}$$

where $\gamma = 0{\cdot}5771$. For an attractive field the cross-section becomes infinite when m is approximately equal to $(n+\frac{3}{4})\,\pi$, where n is integral. When m is small it is easy to see, using the series expansions for the Bessel functions, that the cross-sections for both the attractive and the repulsive field become equal to $\pi m^4 a^{-2}$, the result obtained by the use of Born's formula (21).

Summary.

The quantum theory of collisions is applied to the motion of gas atoms. Using the rigid sphere model, the range of validity of the classical theory of free paths, viscosity, and diffusion is determined. The use of quantum mechanics greatly improves the applicability of this model to the viscosity of helium. The scattering of atoms is considered in detail, and the possibility of experimental proof of the Bose-Einstein statistics is discussed.

Emission of Metallic Ions from Oxide Surfaces. I.—Identification of the Ions by Mobility Measurements.

By Luang Brata, Wills Physical Laboratory, University of Bristol.

(Communicated by A P Chattock, F.R S.—Received March 7, 1933.)

Introduction.

In two recent papers* an account was given of the determination of the mobility of ions of the alkali metals in various gases. It was shown that in argon, krypton, xenon and nitrogen the experimental results could be represented with considerable accuracy by the relation

$$K = \frac{A}{\sqrt{\rho(D-1)}} \cdot \left(1 + \frac{m}{M}\right)^{\frac{1}{2}} \text{ cm./sec./unit electrostatic field,} \qquad (1)$$

where K is the speed of an ion of mass M in a gas of molecular weight m, density ρ and dielectric constant D, at a pressure of 760 mm. of mercury and a temperature of 20° C. For a given gas A is a constant; it changes slightly from one gas to another, the extreme values being 0·48 for nitrogen and 0·56 for xenon. The equation is very similar to one deduced by Langevin using a simple atomic model. In his equation A depends upon the "size" and polarizability of the gas molecules, but for the particular case of a highly polarizable gas it approaches a limiting value of 0·51.

The experimental facts summarized above suggest that in the more polarizable gases the "size" of the ions has little influence on their mobility. This view was confirmed by further experiments in nitrogen which showed that the mobility of even large clustered ions like (Na^+, NH_3) and (Na^+, $2NH_3$) is also given by the same relation which holds for the simple alkali ions.

The first object of the present experiments was to determine if equation (1) expressed the results for ions other than those of the alkalis which are peculiarly simple in possessing a rare gas electronic structure. If the mobility of any ion of mass M, in a given gas, is given by equation (1) using the value of A appropriate to the gas, then the mobility of an ion in this gas gives a measure of its mass. If the mobility can be measured accurately, as it can be with the four gauze method, we have in effect a mass spectrograph of low resolving power.

---

* Tyndall and Powell, 'Proc. Roy. Soc.,' A, vol. 136, p. 145 (1932); Powell and Brata, 'Proc. Roy. Soc.,' A, vol. 138, p. 117 (1932).

This principle is of some importance as it offers a means of identifying positive ions in gases at high pressures. The method can be used for example to study the ionic products of collision processes in the glow discharge in gases to which controlled amounts of impurities have been added. For some purposes, as for example in determining the nature of the positive ions emitted from the sources to be described, the low resolving power is unimportant and the method is convenient in that it is very simple to use and that a complete mass analysis can be obtained very quickly The successful application of the idea of identifying ions by measuring their mobility in a gas depends upon having a method of good resolving power for ions of different mobility and in working in gases of a sufficient degree of purity where the clustering effects of impurity molecules do not obscure the nature of the ions.

Experimental Method.

The choice of ions of gallium, indium and thallium for these experiments was determined by the following considerations In previous experiments it has been found* that a Kunsman source evaporates large numbers of neutral as well as charged atoms. This was proved as follows. Two sources, one of cæsium and one of sodium ions, were mounted side by side , the sodium source was " aged " until it gave only sodium ions, *i e.*, until all traces of more easily ionized alkali impurities had been driven off ; the cæsium source was then heated for a quarter of an hour. On again examining the ions emitted from the sodium source it was found to emit only cæsium ions, and this " induced " emission persisted for some time.

The actual number of cæsium ions emitted by this sodium source was rather large, in the sense that if the superficial area of the source were assumed to be its true area, the cæsium emission corresponded to the evaporation of a layer of the metal a hundred atoms or more thick. Further, the cæsium emission " induced " in the sodium source seemed to be roughly proportional to the time of heating the cæsium source at a constant temperature. Now the alkalis evaporate entirely as neutral atoms or molecules from the pure metals, and a film a hundred atoms thick will, from the point of view of positive ion emission, have the properties of the pure metal. It seemed necessary therefore to assume that the emission from the source was from an effective area very much larger than the superficial area. The simplest assumption was that the cæsium was distributed as an ionized layer not more than one atom thick over a large effective surface. We return later to a discussion of the significance of these results in the theory of the emission from Kunsman sources. The behaviour of the cæsium source which is typical of the alkalis in general suggested that it might be possible to make satisfactory long lived sources of positive ions by direct evaporation of various metals on to an iron oxide surface; that if the ionization potential of the metal were sufficiently low, the oxide, when heated in turn, would emit the required positive ions. Experiments were therefore made with gallium, indium and thallium, which have ionization potentials of 5·97, 5·76 and 6·08 volts respectively, with the apparatus shown in fig. 1.

---

* Tyndall and Powell, *loc. cit.*

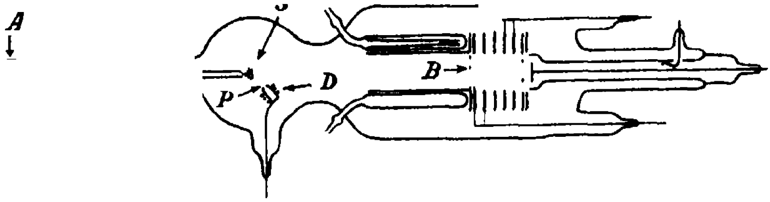


Fig. 1.

The molybdenum pot, P, containing one of the metals, could be heated by electron bombardment from the filament D. The evaporating metal condensed on the source S, which consisted of a few turns of tungsten wire in the form of a cone filled with iron oxide, finely ground and bound together by heating. After the evaporation, the source was moved up by means of a magnet operating on the iron in the slide A, until it nearly touched the copper grid, B, of the analysing part of the apparatus. The apparatus was then filled with nitrogen at a suitable pressure (6 mm.), the source was slowly heated until a current of about half a microampere flowed from it to the plate, B, and the mobility spectrum of the positive ions was determined.* The methods employed have been adequately described in previous papers.

One of the early curves obtained in this way using thallium is shown in fig. 2. Three different groups of ions are present in both the second and third orders. The low mobility peak, which will be shown to be due to singly charged thallium, is predominant, but it is accompanied by peaks at frequencies which identified them as due to sodium and potassium ions. The sodium and potassium were found to be present as an impurity in the ferric oxide which formed the base of the source. Although probably present in small quantities compared with the thallium the ionization potential of these metals is low and their ions are emitted preferentially. If we accept the picture of the emission process already given this merely means that the positive ion work function is lower for the alkalis than for the other metals. Langmuir and Kingdon,* and Killian† show that a higher temperature is required to evaporate monatomic layers of potassium than of rubidium ions and of rubidium than of cæsium ions, from surfaces of pure tungsten.

---

* Powell and Brata, *loc. cit.*

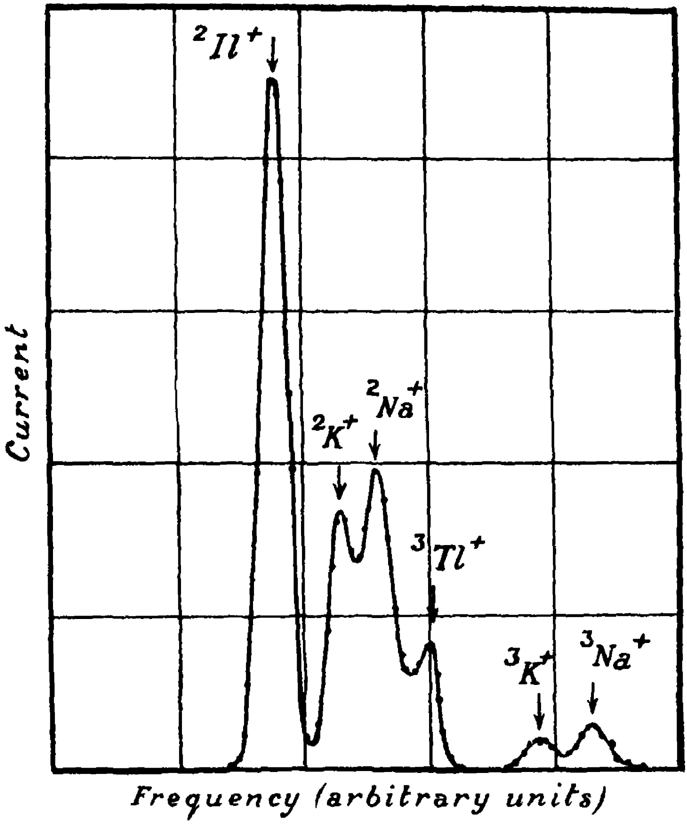


FIG. 2.—Nitrogen.

If the alkali impurities are removed by long heating of the oxide and if the first sample of metal evaporated from the pot, which contains a larger proportion of alkali than that which is evaporated subsequently, is rejected, an almost pure source of thallium ions can be prepared. If indium is substituted for thallium similarly satisfactory sources can be produced. Fig. 3 shows the results obtained with an indium source in argon.

---

[*] 'Proc. Roy. Soc.,' A, vol. 107, p. 61 (1923).

[†] 'Phys. Rev.,' vol. 27 p. 578 (1926).

Identification of the Ions.

The identification of the thallium and indium ions produced in this way depends upon the determination of the mobility of the ions in nitrogen. The mobility of the alkali ions previously measured, is plotted against their mass in fig. 4. It will be seen that the results for indium and thallium lie precisely on the line drawn through the points for the alkalis. The same applies to the result for gallium which was obtained by the same method. There is thus no doubt that we are actually producing the required ions. Further, the results confirm the conclusions of the previous papers that in nitrogen the mobilities of different ions are expressed with remarkable accuracy by equation (1) irrespective of the "size" of the ions.

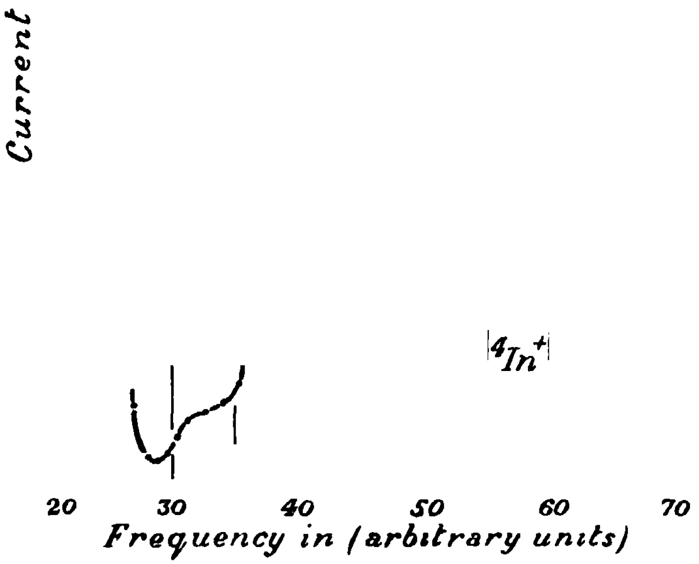


FIG. 3.—Argon.

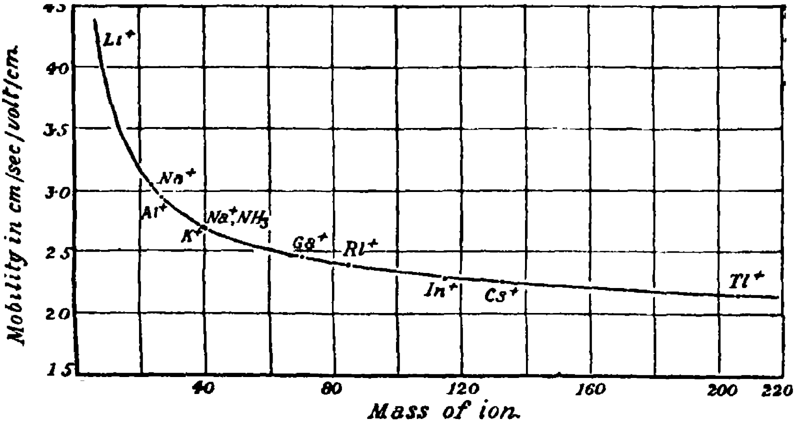


FIG. 4.—Nitrogen.

The Mobility of the Ions in the Rare Gases.

The mobilities of thallium and indium ions have also been determined in the gases argon, neon and helium. The results are shown in Table I, and plotted in fig. 5. It will be seen that in helium the results lie above the line drawn through the points representing the experimental results with alkali ions, in neon they lie practically on the corresponding curve and in argon below it. It is evident that, unlike nitrogen, argon is not a satisfactory gas with which to disperse the ions, since the determination of the mobility does not lead to an unambiguous determination of the mass. This is surprising in view of the fact that the results for the alkali ions in argon can be represented by an equation of type (1), which seems to suggest that the important force between the ions and the gas atoms during gas-kinetic collision is that due to the polarization of the gas atoms and that "size" effects are of no importance in this gas. That the result is real and not due to the presence of impurities in the argon is proved by the fact that on many curves, as for example in fig. 2, peaks due to sodium and potassium ions are present and the mobility of these

Table I.—Mobility of Ions in Gases. Values in cm./sec./volt/cm. at N.T.P.

Ion.	Nitrogen.	Argon.	Neon.	Helium.
Gallium	2·43	—	—	—
Indium	2·29	2·14	7·08	22 0
Thallium	2 17	2 05	6·63	20·3

ions corresponds with the values previously determined. Before this failure of argon was realized an attempt was made to use it as a gas to disperse the ions

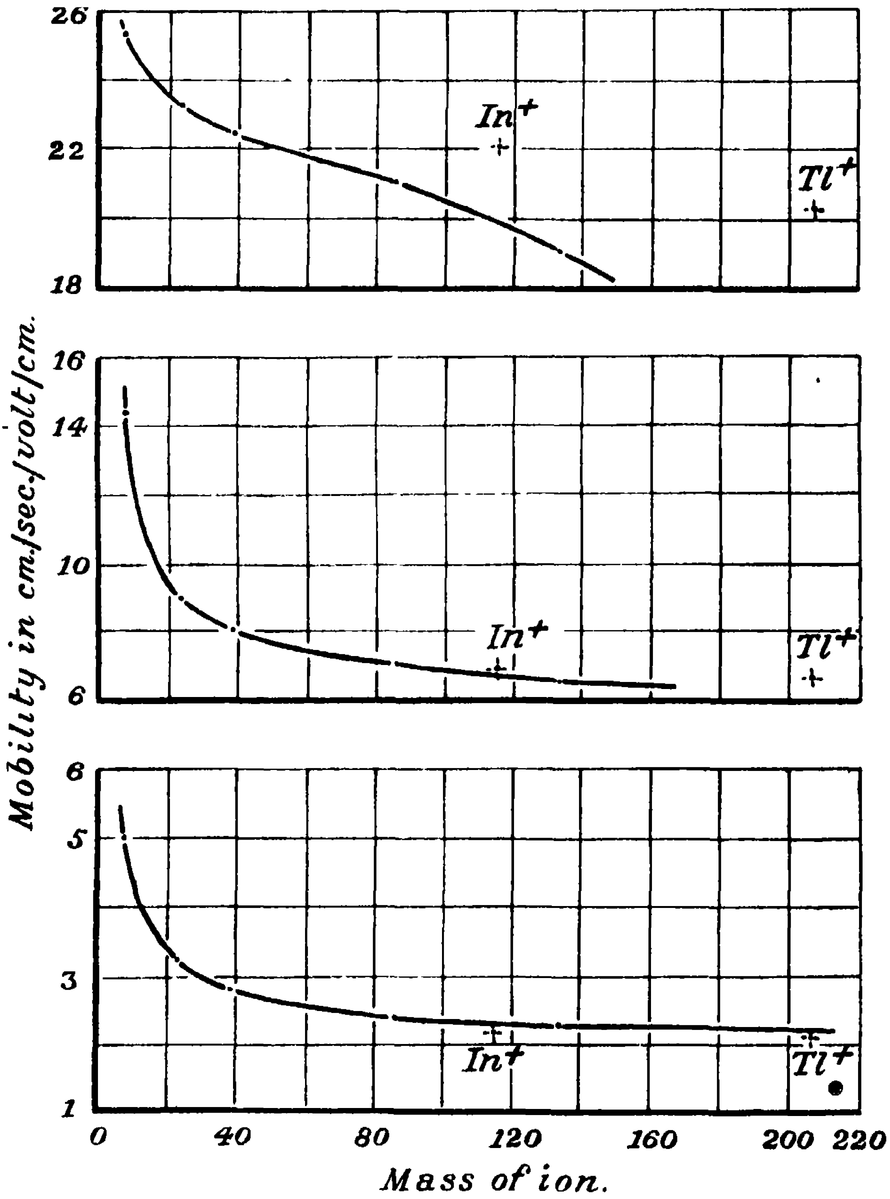


FIG. 5.—Above, helium; middle, neon; below, argon.

emitted from Kunsman sources of the alkaline earths. The results were confusing and no definite identification was made. The reason for this is now clear. The change of mobility due to the "size" of the ions is not serious in argon and when allowances are made for it, it is possible to identify the ions, produced in a glow discharge in the gas, by determinations of their mobility. With helium the difficulties are more serious.

It is difficult to explain the results in helium that the heavy indium and thallium ions move faster, for example, than sodium and potassium. The result may be associated with secondary effects such as the forces produced by the polarization of the ion by the induced dipole on the gas atoms.

Life of the Sources.

A thallium or indium source prepared by the method described will in general give a current of one microampere for several hours. A typical curve showing the decay of the current with time is shown in fig. 6. At the beginning of the run 90% of the positive emission was carried by thallium ions and at the end of 85%. The total quantity of electricity taken from the source during the heating was $7 \cdot 5 \times 10^{-3}$ Coulombs of which $6 \cdot 5 \times 10^{-3}$ Coulombs may be assumed to have been carried by thallium. This corresponds to the transference of $1 \cdot 5 \times 10^{-5}$ gm. of thallium from the source. If we assume that the thallium is on the oxide surface as a film of area equal to the superficial area of the source, and further that every thallium atom evaporates as an ion, then the initial thickness of the film at the beginning of the experiment must have been equal to 3×10^{-5} cm.—a layer a thousand atoms thick.

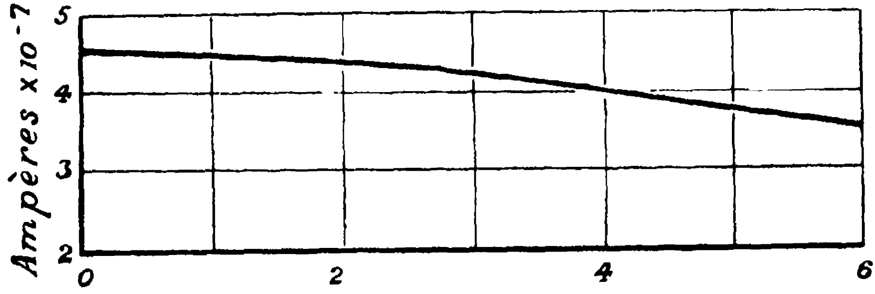


FIG 6.

The indium sources behave similarly to those of thallium. With gallium, however, the positive ion emission falls off very rapidly with time. The only curves taken with gallium sources have been made in nitrogen gas, and in no experiment has a source lasted longer than 30 minutes.

I am greatly indebted to Professor A. M. Tyndall for his interest and valuable advice, and to Dr. C. F. Powell for his assistance during the course of the experiments, and in preparing the work for press. The work was made possible by the grant of a Fellowship by the Rockefeller Foundation of New York.

Summary.

(1) A method of producing sources of positive ions of thallium, indium and gallium is described.

(2) The ions are identified by determining their mobility in nitrogen gas. In this gas it is shown that the mobility of an ion gives an unambiguous measure of its mass. The mobility method therefore offers a means of determining the mass of positive ions in a gas at high pressures. The mobility, K, of any positive ion of mass M in nitrogen gas of molecular weight m at 17° C. is given by the equation

$$\mathrm{K} = \frac{0 \cdot 48}{\sqrt{\rho\,(\mathrm{D} - 1)}} \cdot \left(1 + \frac{m}{\mathrm{M}}\right)^{\frac{1}{2}} \text{ cm./sec./unit electrostatic field,}$$

where ρ is the density of the gas and D its dielectric constant.

(3) The mobility of thallium and indium ions has been determined in helium, neon and argon.

Emission of Metallic Ions from Oxide Surfaces. II.—*Mechanism of the Emission.*

By C. F. Powell and Luang Brata, Wills Physical Laboratory, University of Bristol.

(Communicated by A. P Chattock, F R.S —Received March 7, 1933.)

Introduction.

There are two possible explanations of the behaviour of the sources of positive ions described in the previous paper. We can regard them, firstly, as a kind of modified Kunsman source. It is now well established that when atoms of ionization potential V volts strike a surface of work function ϕ volts, the atoms will be ionized if $\phi > V$. For cool surfaces there may be an accumulation of ions on the surface, if the supply of atoms is maintained, with a consequent reduction of the work function, so that after a time any more atoms arriving are no longer ionized. For hotter surfaces the time spent by the ions on the surface becomes smaller and they are driven off in virtue of the energy imparted to them by the thermal impact of the underlying surface atoms. At higher temperatures every atom striking the surface comes off as a positive ion. The essential condition, $\phi > V$, which must be fulfilled if the atoms shall be ionized on impact with the surface has been confirmed by the experiments of Langmuir and Kingdon* with a number of different surfaces. Thus cæsium (I.P. = 3·8 volts) is ionized by hot tungsten for which ϕ equals 4·53 volts, but not by tungsten covered with a monatomic layer of thorium for which $\phi = 2{\cdot}66$ volts. Similarly copper (I.P. = 9·0 volts) is ionized by a tungsten surface covered with a monatomic layer of oxygen atoms, but not by pure tungsten. The most natural explanation, therefore, of the emission from Kunsman sources would seem to be that the ferric oxide supplies a surface of large electronic work function by which the alkali atoms are ionized. The distribution of the alkali throughout the source and its gradual diffusion to the surface through the conglomerate of microcrystals furnishes a convenient means of maintaining the surface layer of atoms. The evaporation of some neutral atoms by the surface may be regarded as a consequence of the accumulation of the alkali in patches. The alkali may, for example, reach the surface by migration up the cracks between the crystals of which the surface is composed. If the patch is more than one atom thick the work function of the surface may, as in the experiments of Langmuir to which we have referred, be so reduced that atoms and not ions are evaporated from it.

---

* 'Proc. Roy. Soc.,' A, vol. 107, p. 61 (1925).

In the experiments which have already been described in which the alkali from one Kunsman source is evaporated on to a second, it was shown that the simplest explanation of the large numbers of ions emitted was that the real surface area of the source is many times greater than the superficial area. We can make the same assumption to explain the behaviour of the new sources. The source, therefore, owes its long life, in this view, to this very large effective area provided by the ferric oxide. By contrast a surface of a pure metal of the same superficial area as the source, covered with a complete layer of ions one deep would yield a current of 1 microamp for less than 6 seconds.

A second less probable explanation of the mechanism of the emission is that with indium and thallium some of the atoms evaporating from the pure metal come off as ions. At the temperatures at which the emission takes place the metals in bulk are molten, and Goertz* has shown that the electronic work function of some metals rises rapidly beyond the melting point. An experiment was therefore made in which the positive ions emitted from a heated platinum boat filled with thallium were examined by mobility measurements in nitrogen. Only sodium and potassium ions were detected during a run of 1 hour which resulted in the evaporation of 30 milligrams of thallium. This second explanation must therefore be abandoned.

In order to investigate the behaviour of these sources in more detail experiments were made in which a known amount of one of the metals could be deposited on to an iron oxide base which had previously been heated for a long period to remove traces of alkali impurities. A preliminary account of these experiments has been published elsewhere.† A diagram of the apparatus is shown in fig. 1. The furnace B, which contains the metal to be deposited, can be heated by an internal tungsten filament. With a given heating current, i_B, through the furnace, a constant beam of neutral atoms issues through the hole in the top of the furnace towards the centre of the bulb. The density of the atom stream at the centre of the bulb corresponding to any value of the heating current, i_B, can be determined by means of the tungsten strip, F. Thallium and the other metals used in these experiments are ionized by impact with a tungsten surface, covered with a monatomic layer of oxygen atoms, and maintained at suitable temperature. If the tungsten strip F is in this condition every atom striking it is ionized and can be dragged to the plate P and the positive ion current, i_p, to P, measured. The strip is flashed at a high temperature and allowed to cool to about 1000° C. At this temperature, if the apparatus is not too rigorously degassed, the strip gradually picks up a monatomic layer of oxygen atoms so that the electronic work function of the surface rises to 9·2 volts. We can follow the formation of this oxygen layer by bombarding the filament with a neutral beam of thallium atoms. The filament is flashed, allowed to cool to 1000° C. and the positive ion current plotted against the time. A typical curve obtained in this way is shown in fig. 2, *a*. The positive ion current becomes saturated after about 20 minutes in this particular instance and we can regard the filament as completely covered with its oxygen layer after half an hour.

---

* 'Phys. Z.,' vol. 24, p. 377 (1923); vol. 26, p. 206 (1925), 'Z. Physik,' vol. 42, p. 329 (1927); vol. 43, p. 531 (1927).

† Powell and Brata, 'Nature,' vol. 131, No. 3301, p. 168 (1933).

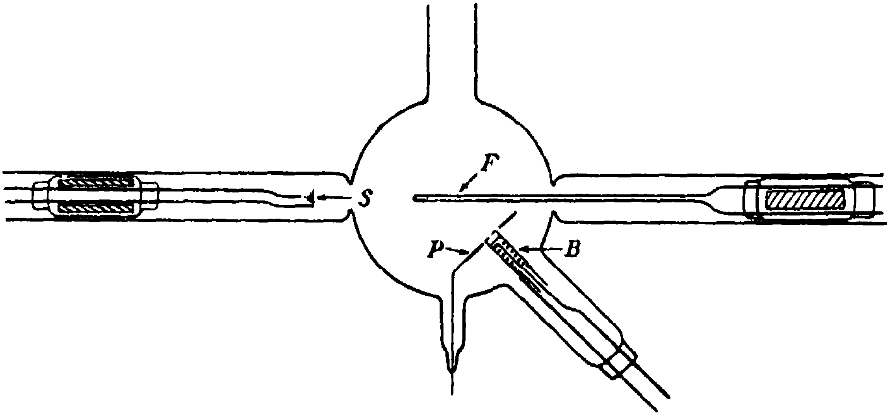


FIG. 1.

With a given stream of neutral atoms from the furnace, the positive ion current, i_p, varies with the filament current, i_f (*i.e.*, with the filament temperature) in the way shown in fig. 2, *b*. At low temperatures the thallium ions are not evaporated by the filament and the atoms reaching it either accumulate on the surface or re-evaporate from it uncharged. At high temperatures the oxygen film itself begins to evaporate. The very limited region over which the positive ion current is at a constant maximum is associated with the shortness of the tungsten strip used in this apparatus. The temperature gradient down the strip is large and only a limited portion of the strip is at a temperature within the critical range.

The positive ion current from the filament to the plate P saturates at voltages well below those at which we should expect the positive ions to produce a serious number of secondary electrons on impact with the plate, fig. 2, *c*.

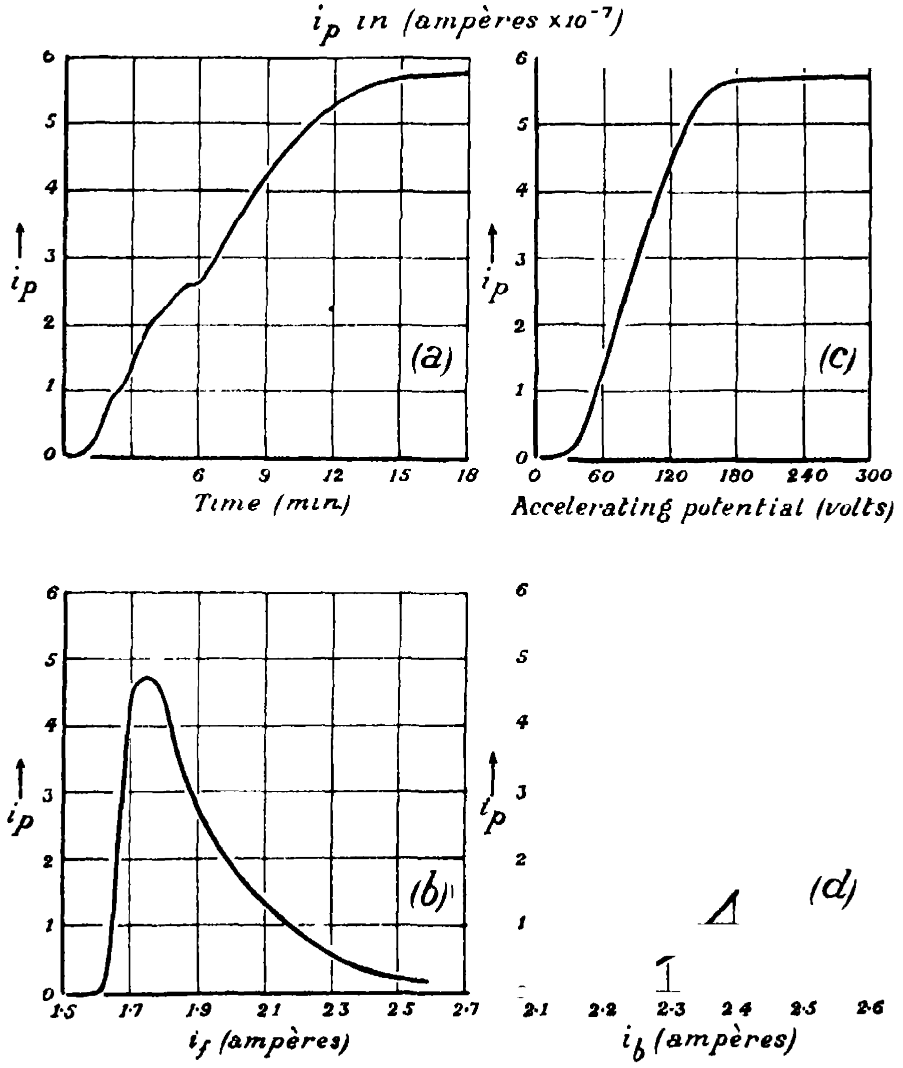


Fig. 2.—(*a*) i_B = 2·63 amps., i_f = 1·75 amps. (*b*) i_B = 2·58 amps (*c*) i_B = 2·63 amps , i_f = 1·75 amps. (*d*) i_f = 1·75 amps.

The positive ion current from the tungsten strip, when i_f equals 1·75 amperes, fig. 2, *b*, gives therefore a measure of the number of neutral atoms hitting the active part of the strip. The area of this active portion of the strip can only be estimated so that the actual stream density of atoms in any particular experiment cannot be found precisely. The relative intensities for different furnace current can, however, be found accurately. The positive ion current for different furnace currents is shown in fig. 2, *d*.

Deposition Experiments.

When the calibration measurements on the furnace have been made, known amounts of a given metal can be deposited on the oxide source if it is moved in turn on its slide into the centre of the bulb. The actual number of atoms deposited in any experiment is known approximately, and the relative numbers in different experiments accurately.

After the deposition of the metal, the oxide source is heated by a given current, i_0, so that it comes to a given equilibrium temperature. The positive ion current from it to plate P is measured and plotted against the time from the moment of switching on the heating current i_0. A typical curve is shown in fig. 3, *a*. The area under the curve represents the quantity of electricity taken from the source, *i.e.*, the number of the deposited atoms which are re-emitted as ions. The estimated number of atoms deposited from the furnace is shown by the hatched rectangle. It will be seen that a very large proportion of the deposited atoms, possibly all of them, come off as ions.

A number of conclusions can be drawn from these results. Since the proportion of deposited atoms which are re-emitted as ions is so large we must conclude that very few atoms are reflected by the oxide during the deposition period; further that by the time the temperature of the source has risen to its equilibrium value the atoms must have distributed themselves over the surfaces of the microcrystals composing the source. Otherwise, if the metal were on the surface as a thick layer, a large proportion of the atoms would be evaporated in the neutral state. We cannot say from these experiments whether this migration occurs at room temperatures whilst the metal is being deposited or whether it occurs, very quickly, at higher temperatures, during the short time that the source is rising to its equilibrium temperature.

If experiments are made with progressively larger deposits and if the source is heated to the same temperature after each deposition, curves of the type shown in fig. 3, *b*, *c*, are obtained. It will be seen that in each case the current rises to the same maximum and subsequently falls off more slowly the larger the deposit. The total number of ions emitted is found to be roughly proportional to the amount of metal deposited. We regard this initial maximum of the curves as the characteristic emission at the given temperature, of the outer surface of the source when fully saturated with the metallic ions. This view is supported by the following experiment. The oxide was exposed to

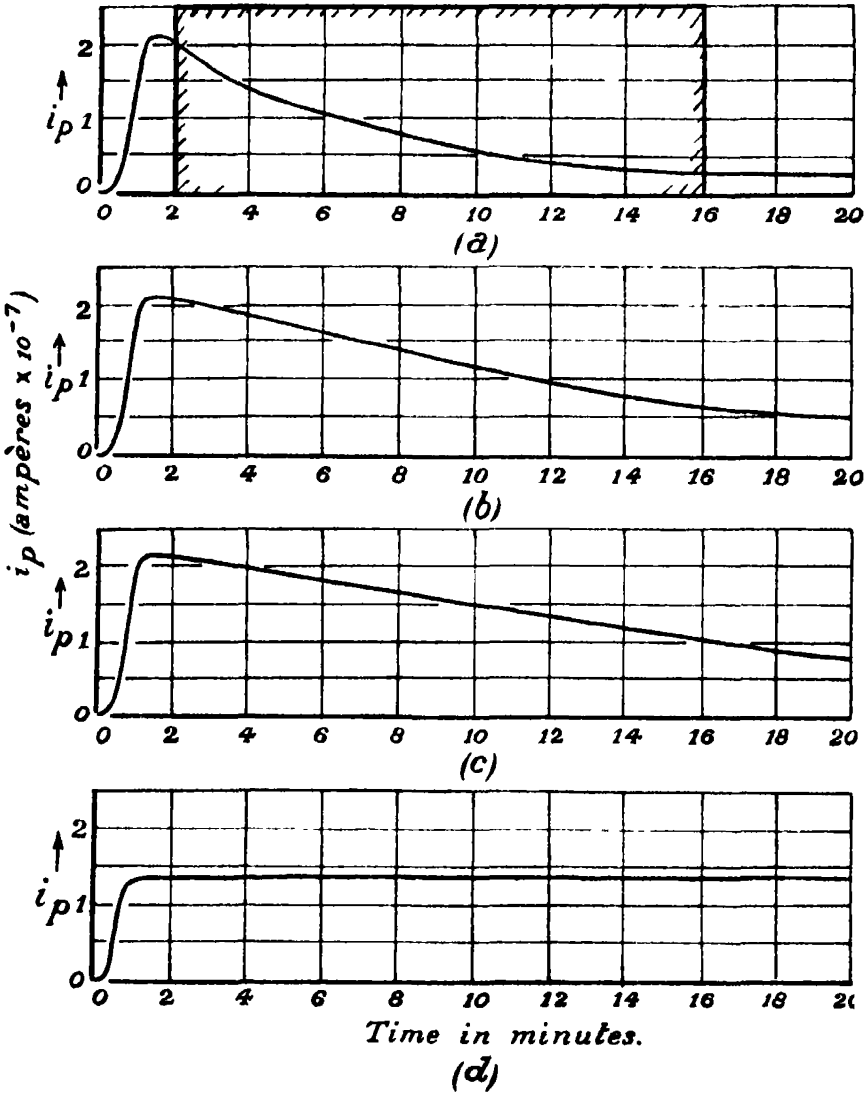


Fig. 3 —(*a*), (*b*), and (*c*) $i_0 = 2 \cdot 8$ amps.; (*d*) $i_0 = 2 \cdot 7$ amps.

a heavy beam of neutral atoms and heated to the equilibrium temperatures corresponding to various currents through its filament; the resulting positive ion emission is plotted against the current heating the oxide, i_0, in fig. 4. It will be seen that the positive ion current corresponding to the equilibrium temperature given by $i_0 = 2 \cdot 80$ amps. is the same as the initial maximum in the curves of fig. 3 which were also taken with $i_0 = 2 \cdot 80$ amps.

Since the ions have been shown to migrate freely over the crystal surface it remains to explain why the outer surface is completely covered with ions when the emission starts. The form of the emission curves suggests that as the emission proceeds the outer surface is supplied with ions from the inner surfaces, and that the greater the initial deposition the more easily is the outer surface maintained in a fully covered state. With large depositions the emission may, for example, remain constant over a long period, see fig. 3, *d*. To explain this tendency of the ions to migrate towards the surface we must assume that they have a larger potential energy on the inner crystal surfaces than on the outer surface, but it is difficult to find an explanation of the origin of this potential

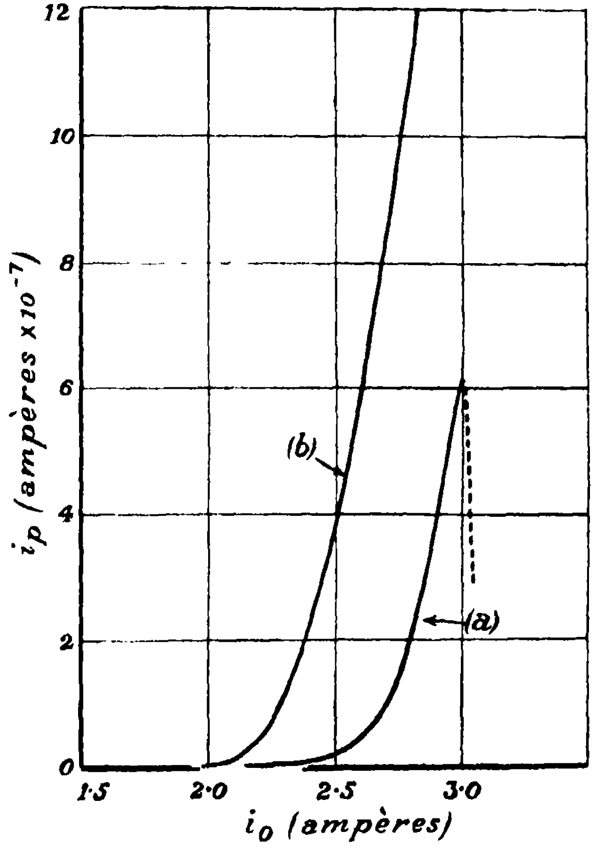


FIG. 4.—(*a*) thallium ; (*b*) indium

difference. If we accept the present picture of the behaviour of the sources it must exist, otherwise the ions would become distributed uniformly over the crystal surfaces and we might expect the initial maximum emission at a given source temperature to be roughly proportional to the number of atoms deposited. We have found other interesting features in the emission curves which call for further experiment. Fig. 5 shows the positive ion emission after a given amount of metal had been deposited on the source. Curve (*a*) was taken after the deposit had been made on a cold source and curve (*b*) when the same deposition was made with the source at a temperature just below that at which positive ion emission is measurable. It will be seen that the two curves are practically identical and that they both show a kink at a time of 7 minutes from the start of the run.

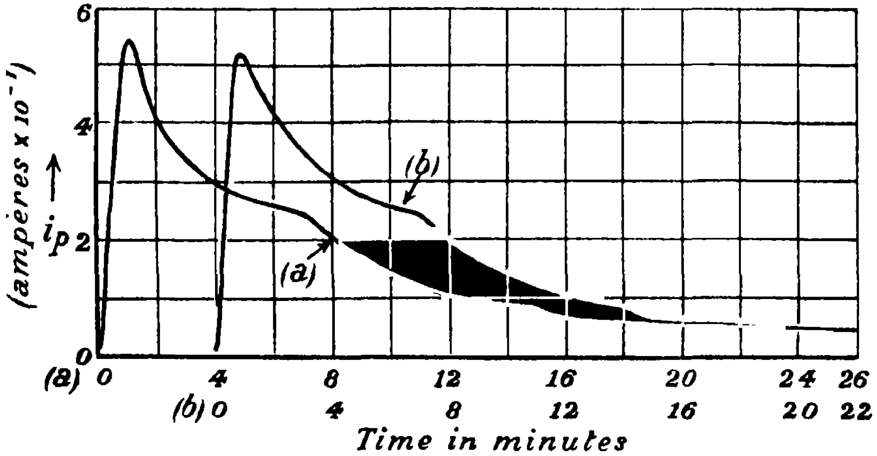


FIG. 5.—(*a*) Deposition hot; (*b*) deposition cold.

Destruction of the Sources by Overheating.

It remains to explain the failure of the gallium sources, which were found to have a very short life. In fig. 4, in which the characteristic emission of the oxide surface when fully covered with thallium ions is plotted against the heating current through the source, it will be seen that the current rises rapidly with increasing temperature, reaches a maximum and then suddenly falls. At higher temperatures the emission becomes erratic, and if the source is maintained at these high temperatures it sinters together to form a very hard, acid resistant mass which can only be put into its old active condition with difficulty. The temperature at which the oxide surface sinters therefore sets a limit to the temperature at which the sources can be heated. With a given kind of ion on the surface, with its own characteristic work function there is a limit to the positive ion current which can be taken from a source of given area. The characteristic emission from a surface covered with indium ions is shown in fig. 4, curve (*b*). It will be seen that the currents at a given temperature are much larger with indium than with thallium, *i.e.*, that the work function for indium ions is less than that for thallium. This lower work function of the indium ions explains the fact that we found it easier to get rid of traces of alkali impurity in working with indium than in experiments with thallium. We have not yet made similar experiments with gallium, but an easy explanation of the failure of the gallium sources presents itself. In order to make a satisfactory analysis of the ions from the gallium source in the mobility experiments we required a certain minimum current. With the area of source there used it was impossible to attain this minimum current, owing to the higher work function of the gallium ions on the oxide surface, without heating the oxide dangerously near the sintering temperature. In most experiments we unknowingly destroyed the gallium sources by overheating.

Experiments with the three metals are now in progress in which the temperature of the oxide sources can be accurately measured; and we may expect to obtain measurements of the positive ion work function for these metals.

It should also be possible to prepare sources of other ions by the present method, provided that the ionization potentials of the metals chosen are less than the electronic work function of the oxide surface. The last quantity is at present unknown. The maximum current obtainable from the sources will, of course, be limited by the sintering temperature of the iron oxide, but it may be possible to overcome this difficulty by using other oxides.

Catalytic Activity of the Kunsman Source.

The Kunsman source and other oxides are widely used as catalysts in the production of ammonia from nitrogen and hydrogen on an industrial scale. The activity of "pure" iron oxide used in this way is much increased by the addition of alkalis. If the iron oxide is overheated its catalytic action is destroyed.

The above facts when taken in conjunction with the results of the present experiments strongly suggest that the catalytic activity of the sources may be due to the presence on the oxide surface of a monatomic layer of alkali ions. The smaller activity of the pure oxide may itself be due to the inevitable presence of traces of alkali as impurity in the oxide. We may assume that the activity of the oxide surface when covered with ions is due to the fact that such a surface can dissociate molecular nitrogen in much the same way that molecular hydrogen is dissociated by a hot tungsten surface. On this view it becomes easy to explain the fact that the catalytic activity is destroyed by certain "poisons," such as arsenic. Suppose, for example, that the oxide surfaces were covered with a monatomic layer of a substance of high ionization potential which remained on the surface as atoms and which could not be driven off at temperatures below the sintering point of the oxide. The electronic work function of the new surface might well be so reduced by the presence of the impurity that alkali atoms would no longer be ionized by it. Such a substance present in comparatively small quantities would completely destroy the catalytic activity of the surface. The same substance would prevent the oxide from acting as a source of positive ions. It should be easy to test these conclusions, and experiments are now in progress using different oxide sources and different poisons.

We are indebted to Professor A. M. Tyndall for his interest in this work and to Professor J. E. Lennard-Jones for a number of discussions we have had with him.

Summary.

(1) An apparatus is described which allows a known number of thallium, indium, or gallium ions to be deposited on to an iron oxide surface.

(2) When the oxide is heated a large proportion, possibly 100%, of the deposited atoms are re-evaporated as ions.

(3) The number of ions leaving the iron oxide surface corresponds in some experiments to the deposition on the superficial area of the oxide of a layer a thousand atoms thick. These atoms must actually have been distributed as a monatomic layer over the large surface presented by the microcrystals composing the source.

(4) If the oxide is heated above about 800° C. it sinters and the original properties of the surface are destroyed. With a given metal on the surface the maximum positive ion current is limited by this maximum available temperature and by the positive ion work function. The failure of the gallium sources in the work described in the previous paper is thus explained.

(5) The experiments suggest an explanation of the characteristics of the positive ion emission from Kunsman sources and of its catalytic activity in the synthesis of ammonia.

The Scattering of Electrons by Metal Vapours. I.—*Cadmium.*

By E. C. Childs, Ph.D., Exhibition of 1851 Senior Student, Clare College, Cambridge, and H. S. W. Massey, Ph.D., Exhibition of 1851 Senior Student, Trinity College, Cambridge.

(Communicated by Lord Rutherford, O.M., F.R.S.—Received March 11, 1933.)

Since the discovery of the diffraction of electrons by gas atoms* a large amount of experimental† and theoretical‡ work has been devoted to the study of electron scattering in gases. As a result it is now possible to recognize the main processes occurring in the collisions with the gas atoms and it has been found that it is usually only necessary to calculate the scattering by the undisturbed field of the atom in order to explain the experimental results. As a consequence considerable simplification is introduced in the theory of the phenomena and it follows that the diffraction effects are mainly determined by the ratio of the wave-length of the incident electrons to the distance from the centre of the atom at which the magnitude of the potential energy of the electron in the atomic field is comparable with its kinetic energy. When this ratio is large (very slow electrons) the angular distribution per unit solid angle of the scattered electrons is independent of angle. As the ratio decreases maxima and minima appear and the diffraction effects become more and more complicated until such electron energies are reached that the ratio begins to increase again. For such energies the simple picture fails but Born's approximation applies and the angular distribution per unit solid angle decreases uniformly with increase of angle. Thus one would expect potassium and argon to give similar angular distributions for electrons with energies considerably greater than the ionization energy of the N electron of potassium but, when the electron energy becomes comparable with this energy, the presence of the outer electron in the alkali metal atom should produce much more complicated angular distributions than are observed for electrons of the same energy scattered by argon.

---

* Bullard and Massey, 'Proc. Roy. Soc.,' A, vol. 130, p. 579 (1931).

† Arnot, 'Proc. Roy. Soc.,' A, vol. 130, p. 655 (1931); vol. 133, p. 615 (1931); 'Nature,' vol. 130, p. 438 (1932); Bullard and Massey, 'Proc. Roy. Soc.,' A, vol. 133, p. 637 (1931); Pearson and Arnquist, 'Phys. Rev.,' vol. 37, p. 970 (1931); Ramsauer and Kollath, 'Ann. Physik,' vol. 12, p. 529 (1932), vol. 12, p. 837 (1932); Hughes and Macmillan, 'Phys. Rev.,' vol. 39, p. 585 (1932); vol. 41, p. 39 (1932); and with Webb, p. 154 (1932); Tate and Palmer, 'Phys. Rev.,' vol. 40, p. 731 (1932); Mohr and Nicoll, 'Proc. Roy. Soc.,' A, vol. 138, pp. 229, 469 (1932); Jordan and Brode, 'Phys. Rev.,' vol. 43, p. 112 (1933).

‡ Massey and Mohr, 'Proc. Roy. Soc.,' A, vol. 132, p. 605 (1931); vol. 136, p. 289 (1932); vol. 139, p. 187 (1933) and in course of publication; also 'Nature,' vol. 130, p. 276 (1932); Allis and Morse, 'Z. Physik,' vol. 70, p. 567 (1931); Feenberg, 'Phys. Rev.,' vol. 40, p. 40 (1932) and vol. 42, p. 17 (1932); Stier, 'Z. Physik,' vol. 76, p. 439 (1932); Henneberg, 'Naturwiss,' vol. 30, p. 561 (1932).

If the above view of the phenomena is correct, it follows that the field of an atom may be approximately determined merely by comparison of the diffraction effects which it produces in scattering electrons with those produced by atoms whose fields are known. All that is necessary is to effect this comparison at a series of different electron energies. A generalization of this method to molecules which have approximately spherically symmetrical fields would also be possible.

In order to make such a method of practical value it is essential to obtain experimental proof of the validity of the method by observations of the angular distributions of electrons of various energies scattered by a number of atoms whose fields are fairly well known. It is also important to determine the conditions under which the complication of electron exchange may be neglected. The present evidence indicates that its effects are negligible for heavy atoms (except possibly at very low velocities of impact) but it is known to be important for the lightest gases, hydrogen and helium, investigated. There are no measurements for atoms of intermediate properties to show how heavy the atom must be in order that exchange may be neglected.

To fulfil these requirements it is clearly necessary to devise methods of investigating electron scattering by metal vapours, as the only monatomic gases are the rare gases with similar electron structure and mercury is the only metal which is sufficiently volatile at ordinary temperature to enable the usual technique to be applied. We have therefore constructed an apparatus for this purpose and have used it to examine the scattering by cadmium vapour. The results obtained, which confirm the general considerations discussed above, leave no doubt that the method employed is capable of application to a considerable number of other metals. A description of the apparatus and results obtained will now be given.

Description of Apparatus.—The principle of the apparatus is illustrated in fig. 1 (*a*). A homogeneous beam of electrons from a gun G is fired through a cloud of cadmium vapour issuing from a circular aperture A of an oven O, and the current due to electrons scattered into the solid angle selected by the collector C measured by an electrometer. The variation of intensity with angle of scattering is measured by rotating the gun about an axis coincident with that of the oven. The distribution of pressure in the jet of vapour issuing from the oven aperture follows a cosine law,* and so the pressure may be taken as constant over the small scattering volume for all angles within the range (25° to 135°) of the experiments. Conditions may therefore be assumed to be the same as for scattering from a uniformly distributed gas.

The detailed construction of the apparatus is illustrated in fig. 1 (*b*). As in previous work by Bullard and Massey (*loc. cit.*) the electron gun G and collecting system C are supported from a large ground joint and can be withdrawn bodily from the scattering chamber for erection and alignment. Both G and C were constructed of eureka sheet. This alloy is very convenient to work and is free from magnetic effects. The gun may be rotated relative to the collector by a smaller ground joint co-axial with the first, the slits defining the electron beam and collected beam being so adjusted that they intersect the axis of rotation. A small copper oven O, supported in a third ground joint at the bottom of the apparatus, has its aperture A also aligned in the axis of rotation

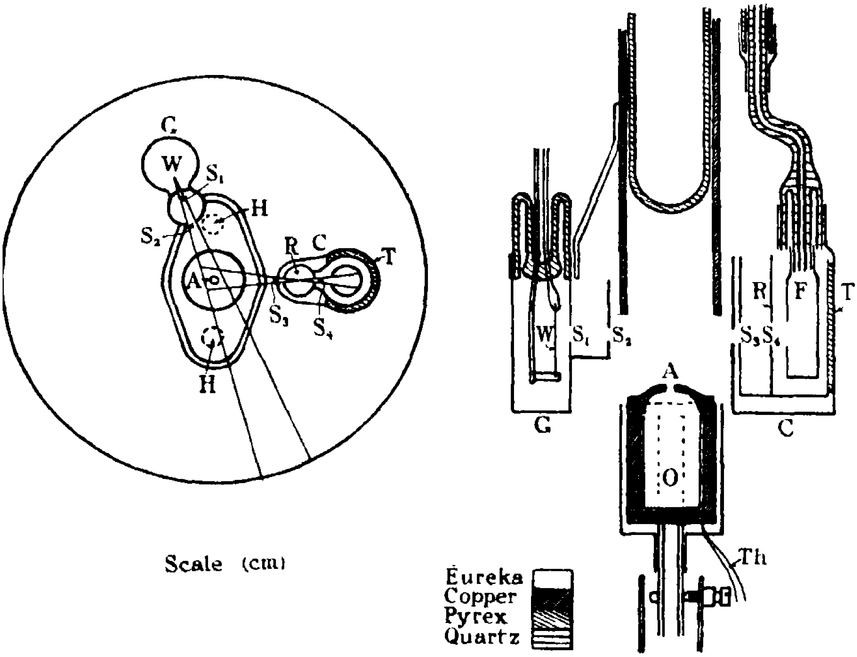


Fig. 1.

(*a*) (*b*)

---

* *Vide* Fraser, "Molecular Rays," Camb. Univ. Press, p. 13 (1931).

of the gun. The scattering chamber is lined with a copper screen in order to maintain a constant potential throughout.

The gun slits S_1, S_2, 6 mm. apart, are respectively 0·15 mm. by 2 mm. and 0·5 mm. by 2 mm., the first defining the beam from the 0·1 mm. diameter thoriated tungsten filament W, the second cutting off stray electrons scattered from the edges of S_1. The filament W is mounted in a pyrex glass pinch clipped into the gun case and the whole is mounted on a copper tube ground to fit a copper plated pyrex tube sealed to the small ground joint. The copper tube extends down as far as possible without interfering with the scattering and serves to collect the bulk of the evaporated cadmium.*

The collecting system consists of the Faraday cylinder surrounded by an ion-repelling electrode R containing slits defining the collected beam. The outer case shields the scattering chamber from the ion-repeller potential. The defining slits S_3, S_4 are 0·5 mm. by 2 mm., 5 mm. apart, all other slits clearing the beam so defined. The Faraday cylinder is supported on the inner part of a quartz internal seal, the outer case, itself supporting the ion-repeller by means of the insulating, tightly fitting pyrex glass tube T, being supported on the outer element of the seal. In this way a long leak path is provided on which cadmium vapour did not condense in sufficient quantity to produce a noticeable leak.

The oven is heated electrically by two heating elements each consisting of 4 cm. of 0·1 mm. diameter molybdenum wire wound in four-bore quartz tubing and inserted in the holes H, fig. 1 (*a*). Loss of heat by radiation is reduced very considerably by the expedient of polishing the oven and surrounding it by a polished radiation shield of eureka The temperature was measured by the eureka-iron thermocouple, Th, and it was found that an expenditure of 10 watts gave an oven temperature of about 450° C.

Experimental Procedure.—In order to reduce the stray scattering to a negligible value the whole apparatus, before the introduction of cadmium, was exhausted to a pressure of less than 10^{-6} mm. Hg, and baked thoroughly for 8 hours with the oven much above the working temperature (generally at about 600° C.). This served to outgas the oven thoroughly. Freshly vacuum distilled cadmium was then rapidly introduced and, as this involved letting down the vacuum, the apparatus was again exhausted and baked for 8 hours at a temperature just below that at which the cadmium began to evaporate freely (about 300° C.). The greatest care was taken to eliminate the presence of mercury vapour in the scattering chamber since cadmium and mercury might be expected to scatter similarly.

---

* Originally it was also intended to investigate the scattering from alkali metal vapours and this design was intended to permit of condensation of the vapour on a liquid-air cooled surface. No measurements have yet been obtained in these vapours owing to surface ionization of the alkali metal vapour at the surfaces of the slits defining the collected beam. This ionization, although giving rise to a very much larger current than that of the scattered electrons, represents the effect of a minute fraction of surface uncovered by alkali metal.

The oven containing the cadmium was then heated gradually to the working temperature, which was made apparent by a sudden rise in scattered current accompanied by a slight fall of oven temperature. A period of about 2 hours was required to do this. Too rapid heating is accompanied by the danger of the cadmium spurting and blocking the oven orifice, thus rendering the laborious preparation abortive. When a satisfactory scattered intensity was obtained the oven temperature was maintained at a steady value by means of control rheostats in the heater circuit.

In order to collect only those electrons which had been scattered without loss of energy a suitable retarding potential (usually about 3 volts less than the full energy of the beam) was applied to the Faraday cylinder. Work was then continued on the measurement of angular distributions until a sudden fall in the scattered intensity indicated exhaustion of the cadmium. The apparatus was then completely dismantled and cleaned preparatory to further measurements.

It is necessary to establish that the observed scattering was due to electrons which have made single collisions with cadmium atoms. The fact that the scattering set in suddenly, accompanied by a slight fall of oven temperature and by marked deposition of the cadmium on the walls of the apparatus, shows clearly that the scattering was due to cadmium. It is not convenient to test for single scattering with cadmium as it is not possible to measure the vapour pressure directly in the scattering space and, owing to the rapid change of slope of the pressure temperature curve, it is unsatisfactory to use change of temperature as a measure of change of pressure. However, if it can be shown that the observed scattering from the cadmium vapour is very much less than that resulting from argon at a pressure which is known to conform to the conditions of single scattering, then there is no doubt that these conditions are also satisfied in the former case. Argon was therefore let into the apparatus through a suitable leak which enabled a pressure of about 10^{-3} mm. Hg to be maintained against the pumps, and the scattering measured, thus establishing the validity of the cadmium results. As the ionization potential and form of the angular distribution curves for this gas are well known, the measurements in argon were also used to determine the contact potential and to check the general working of the apparatus. In this way it was found that the contact potential was 2 volts.

Results and Discussion.—Curves showing the angular distributions of electrons per unit solid angle were obtained for 4, 8, 13, 18, 23, 28, 38, and 48 volt electrons; these are reproduced in figs. 2 and 3. As these curves were obtained on many different days with varying conditions of gun, collector, and oven

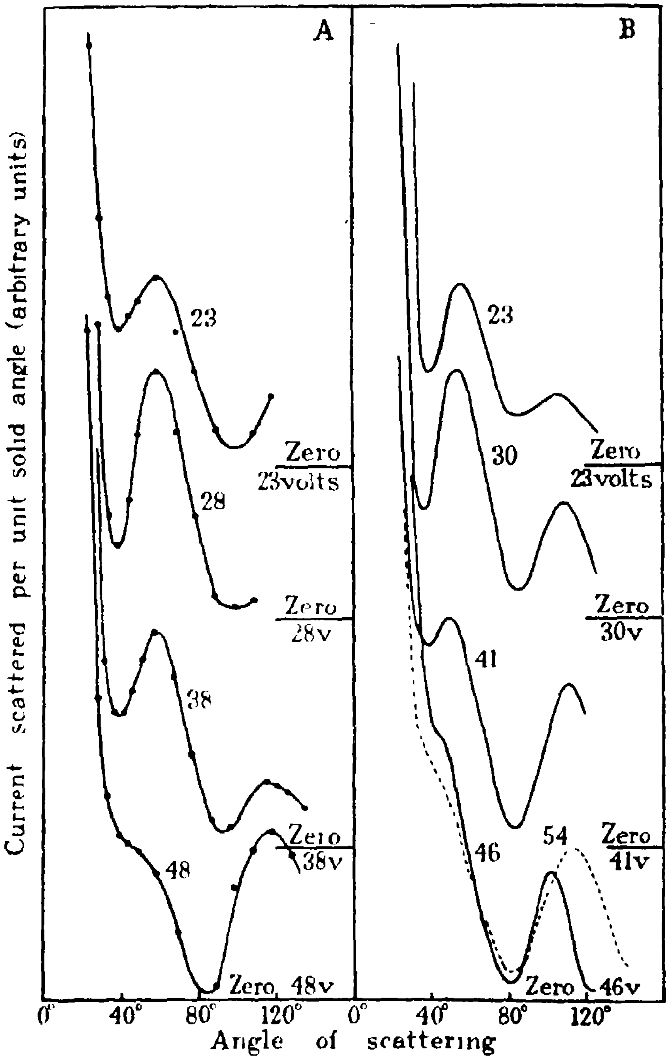


FIG. 2.—Angular distributions of 23, 28, 38, and 48 volt electrons elastically scattered from cadmium vapour. Curves for mercury vapour are given for comparison. • Experimental points. A, cadmium; B, mercury. - - - Jordan and Brode.

alignment, and as slit widths might vary during an experiment owing to deposition of cadmium, a separate curve was obtained with the scattering angle fixed at 60° and the variation of scattered intensity with voltage observed.

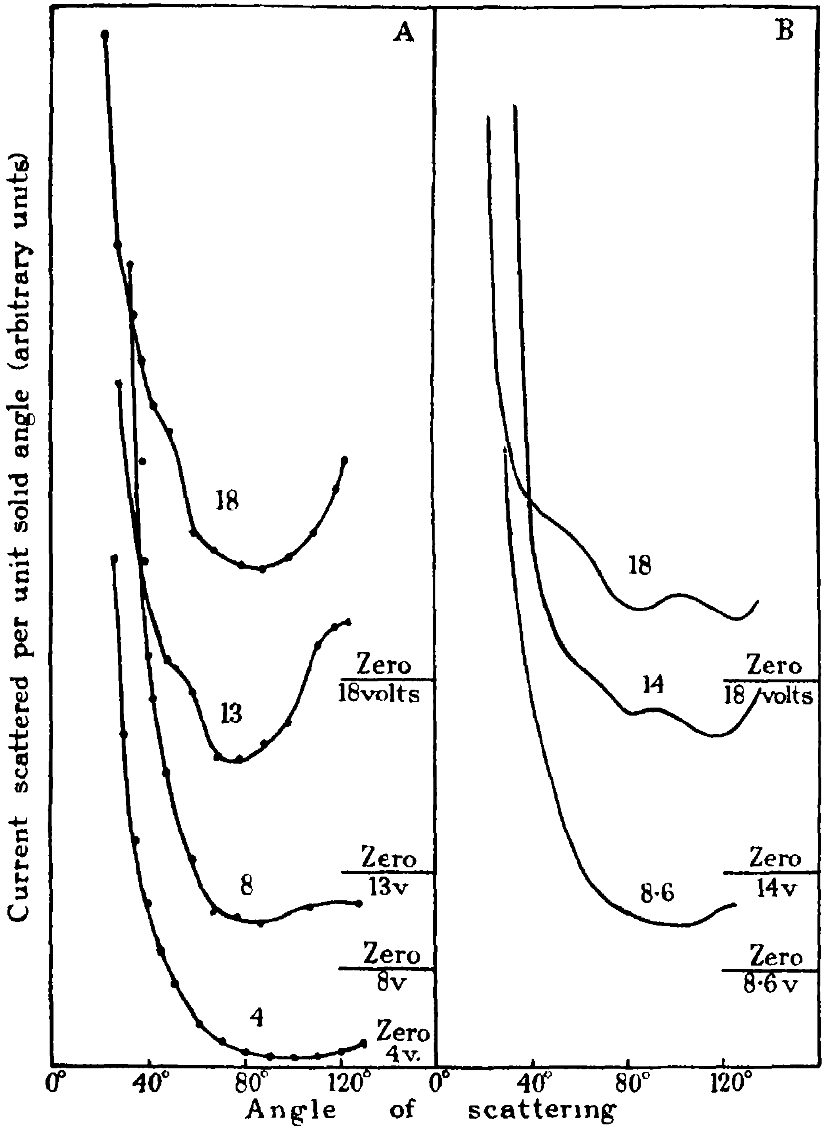


FIG. 3.—Angular distributions of 4, 8, 13, and 18 volt electrons elastically scattered from cadmium vapour. Curves for mercury vapour are given for comparison. ● Experimental points. A, cadmium ; B, mercury.

The results obtained, which are illustrated in fig. 4, enabled the angular distributions to be reduced to the same scale.

By integrating the angular distribution of electrons per unit angle a quantity proportional to the cross-section for elastic scattering between angles of 30° and 130° was deduced as a function of electron energy. The resulting curve is illustrated in fig. 5 and compared with the curve obtained experimentally by Brode* for the total cross-section. Exact agreement is not to be expected as Brode's measurements include the whole angular range and also inelastic collisions. However, it will be seen that the general trend of the curves is the same.

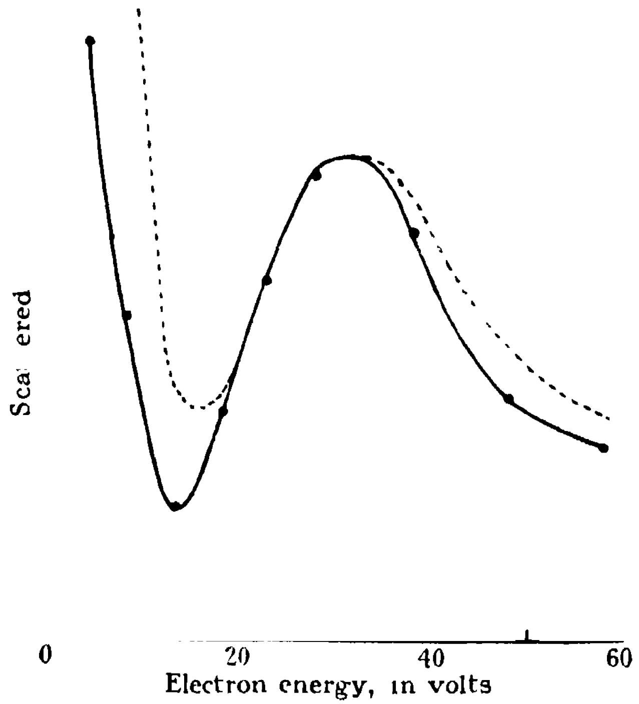


FIG. 4.—Variation of the scattering at 60° with energy of the incident electrons. ● Observed in cadmium; – – – curve for mercury obtained by Arnot.

For purposes of comparison the corresponding scattering curves for mercury vapour obtained by Arnot (*loc. cit.*) and by Jordan and Brode (*loc. cit.*) are also given in figs. 2 and 3. The comparison was effected by adjusting the scale so that the scattered current was the same for both cadmium and mercury at 60° for 30 volt electrons.

---

* 'Phys. Rev.,' vol. 35, p. 504 (1930).

The striking similarity of the two sets of curves is evident. This similarity is not only one of form but persists also in the relative intensities of the scattering at different voltages. Small differences are apparent at large angles, the second maximum for cadmium occurring at a greater angle than for mercury. This difference becomes more marked as the voltage decreases from 48 to 13 volts.

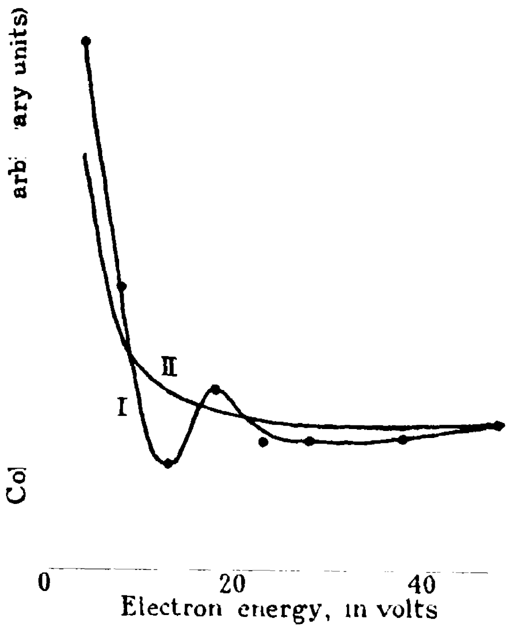


Fig. 5.—Comparison of collision cross-sections for cadmium. I, Cross-section for scattering between 30° and 130° obtained by integration of angular distributions. II, Total cross-section measured by Brode. The scale is adjusted so that both curves agree at 9 volts.

We see, then, that the similar behaviour of mercury and cadmium towards slow electrons, already observed by Brode (*loc. cit.*) in their effective collision cross-sections, is of a very detailed nature. This similarity may at first sight seem surprising in view of the great difference in atomic number between mercury (80) and cadmium (48). However, if we examine the possibilities from the point of view outlined in the introduction we find that this result is to be expected. In Table I the electronic structures of cadmium and mercury are compared. Besides the number of electrons in each shell, the radius of the shell (defined as the distance from the centre at which the charge density of the shell electrons is a maximum) and the ionization potential of the shell are also given. The values have been obtained by means of rules due to Slater.*

Table I.

Shell.		Number of electrons.		Radius (in A).		Ionization potential in volts.	
Hg	Cd	Hg	Cd	Hg	Cd	Hg	Cd
P	O		2	2·15	1·97	10·4	8·95
O	N	18	18	0·50	0·46	250	250
N	M	32	18	0·17	0·15	1800	1500
M	L	18	8	0·074	0·048	6000	7000
L	K	8	2	0·021	0·011	20000	31000
K		2	—	0·006	—	87000	—

To obtain some idea of which shells are effective in scattering electrons of a definite energy the ionization energy is a useful guide, for electrons which have energy much greater than this will not be affected appreciably by the shell and those which have much less will be deviated by the outer shells before reaching the one in question. In this way we see that, for the electron energies used in the experiments described above, the P and O shells of mercury and the O and N shells of cadmium will alone be effective. Referring again to Table I we see that these shells not only contain the same number of electrons for both atoms but that the respective radii are also very nearly the same. Further, the variation with distance of the field due to these respective shells will also be similar as this variation is determined by the radii and ionization potentials concerned. As a consequence it follows that the external fields of the two atoms will be very nearly equal at distances concerned in the scattering. As the diffraction effects depend mainly on the ratio of the wave-length of the electron to the distance at which its kinetic energy is equal to the magnitude of its potential energy in the scattering field we therefore expect the great similarity of behaviour of the two atoms in this respect. From Table I we see that this similarity should persist until the energy of the incident electron becomes comparable with 1500 electron volts, the ionization potential of the M shell of cadmium. The marked difference in the radii and constitution of the inner shells of the two atoms should then become apparent.

---

* 'Phys. Rev.,' vol. 36, p. 57 (1930). These rules do not give very accurate results for the two outer shells of such heavy atoms as mercury and cadmium. Thus they do not give the lowest ionization potentials of the two atoms in the correct order, but the differences between the two are small in any case. The experimental values are given in the table. The radii of the outer shells are probably in the inverse order to that given. However, for the present purpose Slater's rules are sufficiently accurate.

This example illustrates the possibilities of the electron scattering method of investigation of molecular fields. It is clear that further experiments will be very valuable in increasing the power of the method and it is hoped to apply the apparatus described in this paper to the investigation of a number of other metals.

In conclusion we wish to express our appreciation of the interest that Lord Rutherford and Dr. J. Chadwick have taken in these experiments. Our thanks are also due to Dr. Broadway for his advice on the design of the oven.

Summary.

The angular distributions of slow electrons scattered from cadmium vapour have been investigated over the angular range from 25° to 130°. Scattering curves have been obtained for 4, 8, 13, 18, 23, 28, 38, and 48 volt electrons. These curves exhibit maxima and minima closely resembling those appearing in corresponding curves for mercury vapour The significance of these results is discussed and their bearing on the possible application of scattering measurements to the determination of atomic and molecular fields considered.

The Influence of Pressure on the Spontaneous Ignition of Inflammable Gas-Air Mixtures. I.—Butane-Air Mixtures.

By Donald T. A. Townend, D.Sc., and M. R. Mandlekar, B.Sc., High Pressure Gas Research Laboratories of the Imperial College of Science and Technology, London.

(Communicated by W A. Bone, F.R.S.—Received April 8, 1933.)

Introduction.

As part of an investigation into explosions of various hydrocarbon-air media at elevated temperatures and pressures, we have recently been determining the influence of varying initial pressures up to 15 atmospheres on their reactivities during slow combustion up to their ultimate ignition points. And as the results obtained in the experiments on the spontoneous ignition of butane-air mixtures have presented some very striking new features, which seem of undoubted importance in regard to the problem of "knock" in internal combustion engines, we are submitting them in the present communication

Hitherto, few investigators have determined spontaneous ignition temperatures under pressure and, in particular, little is known concerning the behaviours of mixtures with air of the higher members of the paraffin hydrocarbons. The principal research on this problem has been that of Tizard and Pye,* who, employing the adiabatic compression method, found with pentane-, hexane-, heptane-, and octane-air mixtures that with compression ratios of 6·09 to 1 ignition occurred at temperatures of *circa* 300° C. which (*a*) were dependent upon the observed time-lags between compression and ignition, (*b*) varied but little with mixture composition, and (*c*) were lowered slightly as the paraffin series was ascended

Also, by means of his concentric-tube apparatus Dixon† showed a progressive lowering of the ignition temperature in air of the simpler hydrocarbon methane (for a constant lag of 3 seconds) of from 694° C. to 624° C. as the pressure was raised from atmospheric to 7 atmospheres; and in his more recent adiabatic compression apparatus he found 428° C. for a 7½% methane-air mixture with a compression ratio of 12 to 1.

---

* 'Proc. N.E. Coast Eng.,' vol. 31, p 381 (1921); 'Phil. Mag.,' vol. 44, p. 79 (1922) and vol. 1, p. 1094 (1926).

† Dixon and Higgins, 'Trans. Faraday Soc,' vol. 22, p. 267 (1926).

Elsewhere, Dumanois and Mondain-Monval* have heated progressively for about 45 minutes in a steel bomb certain pentane-air mixtures initially made up at about 6·3 kg. pressure and at room temperature; by following manometrically the pressure rise on heating, ignition was observed to occur at temperatures between 220° and 250° C. And in 1924 Tauss and Schultze,† employing a method whereby the fuel was injected with or without spraying into air heated under pressure, showed without spraying a lowering of the ignition point of petrol-air mixtures from 375° C. to 232° C. as the pressure was raised from 2·2 to 10 atmospheres, the corresponding figures with spraying being somewhat higher. They concluded generally that the effect of pressure was so variable with different fuels that no conclusion could be drawn from values at atmospheric pressure as to the order of their ignition temperatures under pressure.

Perusal of the literature‡ shows that generally speaking ignition temperatures of mixtures of the higher paraffin hydrocarbons with air when determined *at atmospheric pressure* are of the order of 500° C.; *under higher pressure,* however, as for example in the experiments of Tizard and Pye, they are more usually of the order of 300° C. It is shown in this communication how this fall in ignition temperature is not progressive with gradual increase in initial pressure but, assuming butane-air mixtures to be typical of the higher hydrocarbon-air mixtures generally, most of it occurs when some definite critical pressure increment, which varies slightly with the composition of the mixture, has been attained.

Experimental.

Method.—The method employed, which has been adapted to meet the requirements of high pressures, originated with Mallard and Le Chatelier,§ and has been adopted by a number of more recent workers.|| The explosive mixture under investigation is rapidly admitted into an evacuated vessel maintained at some accurately known temperature. At temperatures well below the

---

* 'C. R. Acad. Sci. Paris,' vol. 187, p 892 (1928).

† 'Z. Ver. deut. Ing.,' vol. 68, p. 574 (1924).

‡ *Vide* "Flame and Combustion in Gases," Appendix, Bone and Townend.

§ 'Ann. Min.,' vol. 4, p. 274 (1883).

|| Notably Taffanel and Le Floch, 'C. R. Acad. Sci. Paris,' vol. 156, p. 1544 (1913), vol. 157, pp. 469, 595 and 714 (1913); Mason and Wheeler, 'J. Chem. Soc.,' vol. 121; p. 2079 (1922); vol. 125, p. 1869 (1924); Naylor and Wheeler, 'J. Chem. Soc.,' p. 2456 (1931); Laffitte and Prettre, 'C. R. Acad. Sci. Paris,' vol. 187, p. 763 (1928); vol. 188, p. 1403 (1929); vol. 190, p. 796 (1930); vol. 189, p. 177 (1929); vol. 191, p. 329 (1930); vol. 191, p. 414 (1930).

ignition point slow combustion begins and at atmospheric pressure this can usually be followed quite easily by means of a manometer.* With increasing temperature a point is reached when, after a definite time-interval following admission of the gaseous mixture into the vessel, ignition occurs. For the purpose of these investigations the ignition point is taken as the lowest temperature to which the vessel must be heated in order for ignition ultimately to occur In our experiments the time-lags at the ignition point were usually between 15 and 25 seconds, tending to become less at the lower pressures employed and with mixtures containing an excess of air; also increasing the temperature by some 10° C. usually reduced the time lag to about ½ second. Whether or not ignition occurred was critical to 1° C.

Ignition temperatures determined in this way are by no means physical constants, being dependent upon the capacity, surface-volume ratio and material of construction of the heated vessel and possibly also on other factors. The method is a good one from other points of view, however, and the results obtained by it may be taken as comparable relatively one with another.

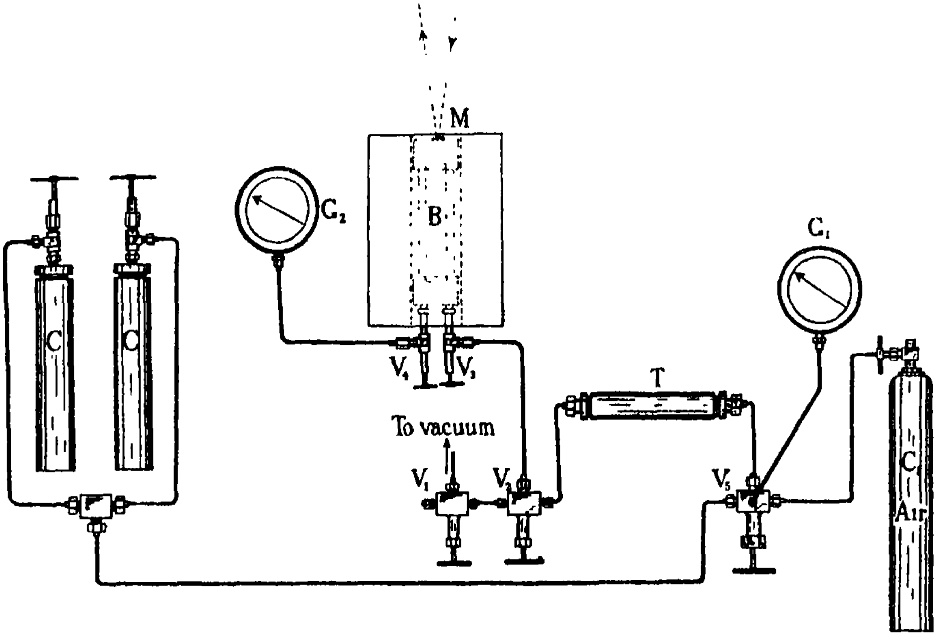


Fig. 1.

Apparatus and Procedure.—The explosion vessel employed, fig. 1, consisted of a nickel-steel bomb, B, fitted for the purpose of the present investigation with a mild steel liner 15 cm. long and 4 cm. in diameter, the capacity being 190 c.c. One end fitting was furnished with inlet and outlet values, V_3 and V_4 and the other with an invar disc recording manometer. The bomb was mounted in a heavily lagged electric furnace, temperature being recorded by a thermo-junction embedded in the bomb wall close to the liner. An interval of at least 20 minutes was always allowed for the establishment of thermal equilibrium before making any determination

---

* Bone and Hill, 'Proc. Roy. Soc.,' A, vol. 129, p. 434 (1930), and Bone and Allum, *ibid.*, vol. 134, p. 578 (1932).

The explosive mixtures were made up and stored under pressures usually between 20 and 30 atmospheres in two thick-walled cylinders C, C maintained at about 50° C. Before each experiment the capacity vessel T was filled from the storage cylinders to some predetermined pressure, so that releasing the gaseous mixture from T into B rapidly filled it to the desired experimental pressure at the trial temperature. The operation of filling the explosion vessel with an explosive mixture to the desired pressure was usually complete in between ½ and 1 second.

At pressures above 5 atmospheres ignition was usually registered on pressure-time records obtained by means of the disc manometer M. At pressures below 5 atmospheres it was observed by the kick imparted either to (*a*) a Bourdon gauge G_2 or (*b*) at pressures below 1½ atmospheres, to a mercury manometer. In all determinations the time-lags between the moment of filling and the occurrence of ignition were checked by means of a stop-watch. Between each determination air under pressure was left in the vessel and on its subsequent release the vessel was evacuated with a Hyvac pump for a period of about half an hour before making a further determination.

Gas Mixtures.—The butane employed was taken from cylinders containing the liquefied gas supplied by the Ohio Chemical and Manufacturing Company, Cleveland, U.S.A. It was stated by the manufacturers to have a purity of 99·9 per cent. and this figure has been confirmed by chemical analysis. Each mixture was allowed to stand 18 hours after making before being brought into use, and samples were taken from the storage cylinders both before the commencement and after the completion of a series of determinations in order to ensure that no change in composition had occurred on account of possible stratification.

Results.

In fig. 2 a series of curves has been drawn showing the observed variation of ignition temperatures with mixture composition at pressures of 1, 1½, 1¾, 2, 3, 5, 10 and 15 atmospheres. At atmospheric temperature and pressure the range of inflammable mixtures lies between those having butane contents of approximately 1·7 and 5·7% respectively.* At the higher temperatures and pressures employed this range would be widened in each direction. We have not extended our observations to these limiting compositions in

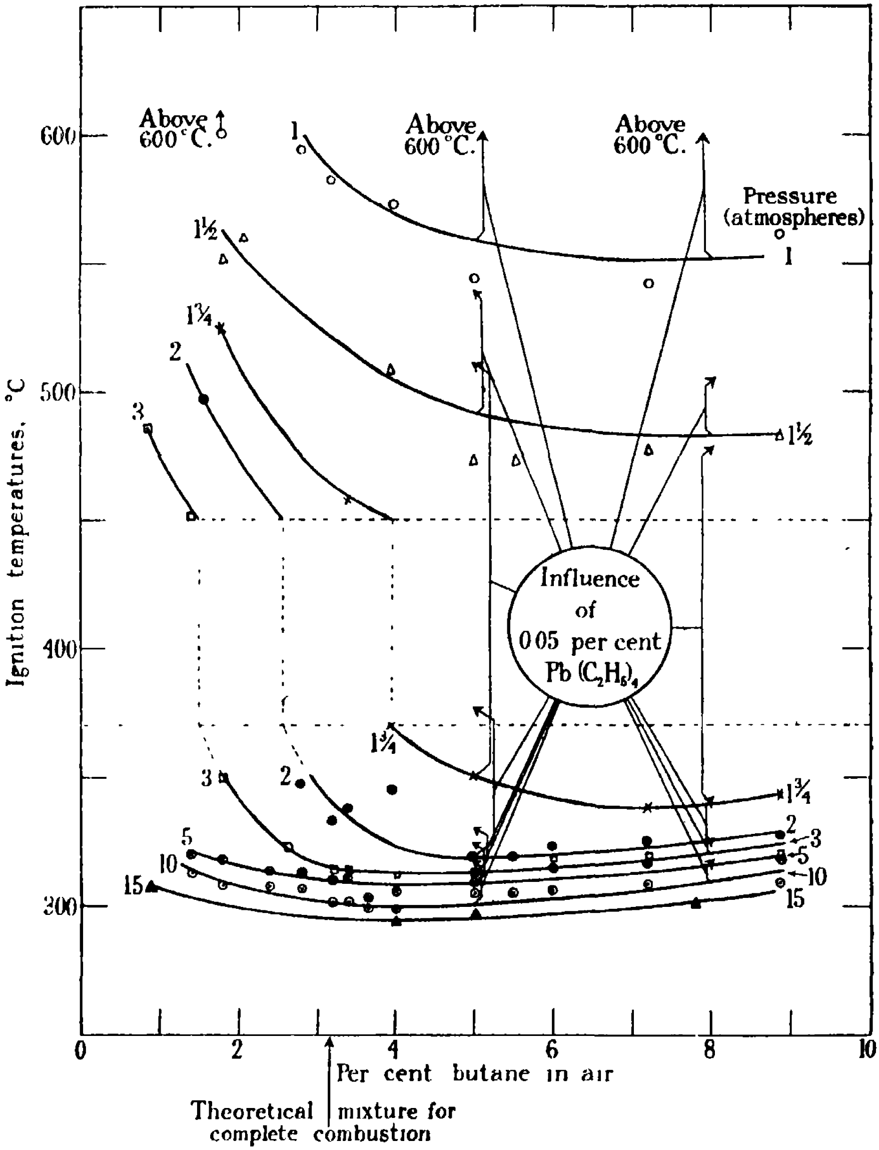


Fig. 2.—○ = 1 atmosphere; △ = 1½ atmospheres; × = 1¾ atmospheres; ● = 2 atmospheres, □ = 3 atmospheres; ⊗ = 5 atmospheres; ⊙ = 10 atmospheres; ▲ = 15 atmospheres.

---

* *Vide* "Flame and Combustion in Gases."

each case, but have confined them to mixtures whose butane contents varied between 0·9 and 8·9%.

It will be observed that, whereas at atmospheric pressure the ignition points gradually fell from above 600° C.* to *circa* 550° C. with increase in butane content in the mixtures, and at 1½ atmospheres pressure from *circa* 560° C. to 480° C., on further increasing the pressure by merely ¼ atmosphere to 1¾ atmospheres a lowering of no less than 130° C to 140° C. in the ignition temperature took place for all mixtures of butane content higher than about 3·5% (the theoretical mixture for complete combustion containing 3·1%). With mixtures containing an excess of air a similar sharp fall in ignition temperature occurred with a definite small pressure increment, but this was now observed at a higher absolute pressure which depended upon the composition of the mixture. At pressures above 5 atmospheres the ignition temperature curves took a new form showing now a slight minimum, which occurred at a mixture composition whose butane content was about 4 per cent.; further increase in initial pressure to 10 and 15 atmospheres merely lowered the curves uniformly a few degrees.

Particular attention is directed to the fact that the ignition temperatures determined in this investigation lie in two well defined groups located above and below the two horizontal dotted lines, fig. 2. Thus at the low pressures (*i.e.*, in almost every mixture below 3 atmospheres) all the determinations lie above 450° C. and at the higher pressures (*i.e.*, in almost every mixture above 3 atmospheres) the determinations lie below 370° C. and usually below 350° C Also the fact that as far as our experiments have proceeded ignition has never occurred at any intermediate temperature between the two groups referred to seems to be of great significance.

The Influence of Pressure.

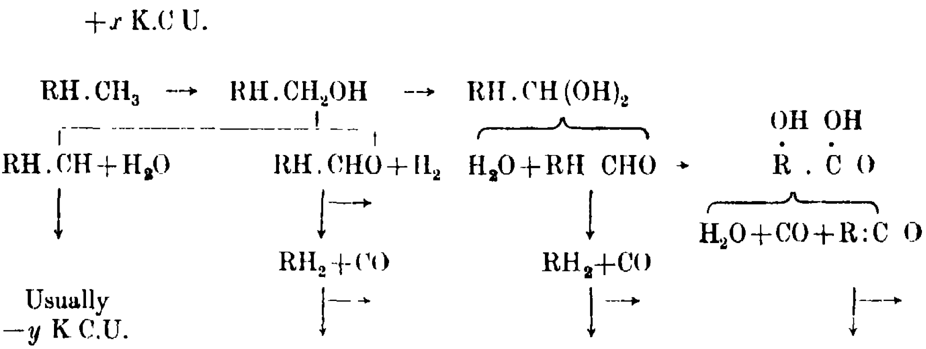
While at the present stage of the investigation it is premature to offer any final explanation of the observed facts we think certain inferences may be made.

In the first place we regard this type of ignition as purely a thermal phenomenon. The pressure-time records indicate that the pressure ultimately developed decreases with increasing time-lag between filling and ignition, as though normal slow combustion had proceeded during the lag period. Also " snap " samples of the explosive mixture taken during this period in circumstances when ignition would ultimately have occurred, while failing to reveal any trace of "peroxides," always showed considerable quantities of aldehydes to be present.

---

* As it was considered undesirable to raise our present apparatus to temperatures higher than 600° C. we have not determined the ignition points at atmospheric pressure of mixtures of butane content below 2·6%.

We think the best explanation of these facts is as follows: according to the hydroxylation view oxidation of a hydrocarbon would be expected to proceed in stages according to some such scheme as fig. 3.



where $RH = C_n H_{2n+1}$

FIG. 3.

The life of intermediate compounds so formed would depend upon temperature, pressure, mixture composition and catalytic influences, effects which, by the recent work of Newitt* and his collaborators in these laboratories, have been strikingly demonstrated to be operative.

At low initial pressures therefore it is not unlikely that in the initial stages of combustion the life of certain intermediate compounds is comparatively short. Also, of the heat evolved in the reactions producing them the whole or part might be absorbed in their thermal decomposition so that little heat might be expected to accumulate in the system. In brief the butane system might rapidly be degraded into one composed of hydrogen, carbon monoxide and lower hydrocarbons, none of which would be expected ultimately to ignite at temperatures below 500° to 600° C.

The influence of pressure would be three-fold, namely, to (i) lower the temperature at which an equivalent chemical reactivity would occur,† (ii) increase the stability of intermediate compounds, which invariably break down with increase in molecular volume,† and (iii) lower the relative heat loss. In these circumstances it is not unlikely that later stages of oxidation of intermediate products would precede their thermal decomposition with consequent further heat evolution, the whole reaction going to completion (ignition) almost immediately Ignition would then occur at temperatures not greatly in excess of that corresponding to the initial slow combustion (*circa* 300° C.) and probably under conditions which at lower pressures give rise only to cool flames.*

---

* 'Proc. Roy. Soc.,' A, vol. 134, p. 592 (1932); vol. 140, p. 426 (1933).

† Newitt, *loc. cit.*

The Function of Antiknocks.

It is well known that antiknocks as, for example, lead tetraethyl, play a prominent role in the slow oxidation of hydrocarbons, *inter alia* usually lowering the temperature coefficients of reaction velocity. Also, as demonstrated by Egerton and Gates† addition of antiknocks to various hydrocarbon-air media is capable of raising appreciably their respective ignition temperatures at atmospheric pressure, as the following figures show :—

Hydrocarbon.	Ignition temperature	Rise in I T. on addition of 0·25 per cent by volume of $PbEt_4$.
	° C.	° C.
Pentane	515	75
Isohexane	525	46
Heptane	430	83
Benzene	(690)	18
Petrol (Shell)	460	82

If the hypothesis now put forward to explain the division of ignition temperatures into two main groups be correct it would suggest that under pressure the influence of antiknocks might in certain circumstances be even greater than at atmospheric pressure ; for at pressures near the critical transition pressure it might be possible to cause a transfer of the ignition points from the lower to the higher group

On testing this supposition it was actually found to be correct, for with two mixtures each containing 5 and 8 per cent. of butane, respectively, and to which 0·05 per cent. of lead tetraethyl had been previously admixed, the following ignition temperatures were determined .—

---

* *Cf* Emeléus, 'J. Chem Soc.,' p. 1733 (1929) and Prettre, 'Bull Soc Chim. Paris,' vol. 51, p. 1132 (1932).

† 'J. Inst. Pet. Tech.,' vol. 13, p. 244 (1927).

Initial pressure.	5 per cent. butane-air mixture.			8 per cent. butane-air mixture.		
	Normal ignition temperature	Ignition temperature with 0·05 per cent. $Pb(C_2H_5)_4$.	Rise.	Normal ignition temperature.	Ignition temperature with 0·05 per cent. $Pb(C_2H_5)_4$.	Rise.
atmospheres	° C.	° C.	° C.	° C	° C.	° C.
1	560	above 600	> 40	562	above 600	> 38
1½	490	539	49	483	502	19
1¾	350	510	**160**	340	478	**138**
2	319	366	47	325	341	16
3	314	330	16	320	325	5
5	308	324	16	316	318	2
10	302	318	16	309	317	8

It will be observed that (*a*) most striking rises of no less than 160° and 138° C. occurred at the pressure (1¾ atmospheres) just adequate to lower the normal ignition temperature from the higher to the lower ignition temperature group (see also fig. 2), also (*b*) when a high enough pressure had been attained, at about 2 or 3 atmospheres, little rise in ignition temperatures occurred their location remaining in the lower group.

Much interest attaches to the detection of the particular intermediate compound or compounds whose survival and further rapid oxidation is responsible for the lower group of ignition temperatures and also to the precise function of antiknocks which may not of necessity be merely one of directly inhibiting oxidation. It is highly probable that pressure plays a predominant part in both these phenomena, and if so a good deal of the work hitherto carried out at atmospheric pressure would fail to throw much light on the matter. On this account we are determining the critical pressure regions for other explosive media and are making a detailed study of the respective influences of the successively formed intermediate products throughout the combustion under pressure.

In conclusion we desire to express our thanks to Mr. L. L. Cohen, M.Sc., for assistance in carrying out some of the experiments as well as to the King Edward Memorial Society, C.P. and Behar, India, for a Fellowship which has enabled one of us (M. R. M.) to devote his whole time to the work.

Summary.

An investigation has been made of the influence of initial pressure on the spontaneous ignition of various butane-air mixtures at pressures up to 15

atmospheres. At atmospheric pressure the ignition points fell from above 600° C. to *circa* 550° C. with increase in butane content in the mixtures and at 1½ atmospheres from *circa* 560° C. to 480° C. Further increase in pressure by only ½ atmosphere caused a lowering of the ignition points by no less than 130° C. to 140° C. for mixtures of butane content higher than *circa* 3·5%; a similar sharp fall occurred with all other mixtures, the critical transition pressure then depending upon the composition of the mixture.

The ignition points were found to lie in two well defined groups the one, for determinations *below* 3 atmospheres, located above 450° C. and the other, for determinations *above* 3 atmospheres, below 370° C. It is suggested that the higher group represents ignition points of the products of thermal decomposition of intermediate oxidation compounds, and the lower group those resulting from ignitions initiated in the primary combustion itself.

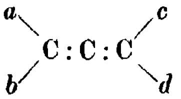
The influence of lead tetraethyl in raising ignition temperatures is very much greater at the critical transition pressures than at atmospheric pressure, an effect which might be anticipated if the above view of the phenomenon be correct.

*An Optically Active Arsonic Acid Possessing Molecular Dissymmetry. Resolution of dl-spirobis-*3 : 5-Dioxan-4 : 4′-di (phenyl-p-arsonic acid).

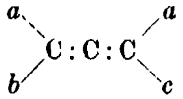
By Charles Stanley Gibson, F.R.S., and Barnett Levin, Chemistry Department, Guy's Hospital Medical School, London.

(Received May 20, 1933.)

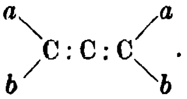
van't Hoff (1875) showed that, according to the principle of the tetrahedral configuration of groups attached to a carbon atom, allene compounds of type I being asymmetric can exist in enantiomorphous forms. It has been pointed out by subsequent authors that molecules of simpler compounds represented by type II are asymmetric, and those of type III are dissymmetric,* and that compounds of all these types should be capable of being resolved into their optically active isomers


$\begin{matrix}a & & c\\ & C:C:C & \\ b & & d\end{matrix}$

I.

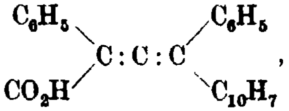

$\begin{matrix}a & & a\\ & C:C:C & \\ b & & c\end{matrix}$

II.

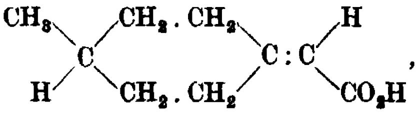

$\begin{matrix}a & & a\\ & C:C:C & \\ b & & b\end{matrix}$.

III.

The first example of an allene compound, whose molecule has an asymmetric configuration, and to be synthesized is αγ-diphenyl-γ-1-naphthylallene-α-carboxylic acid (type II),*


$\begin{matrix}C_6H_5 & & C_6H_5\\ & C:C:C & \\ CO_2H & & C_{10}H_7\end{matrix}$,

but this compound has not been resolved into its optically active components. Perkin and Pope† showed that such a compound as 1-methy*cyclo*hexylidene-4-acetic acid,


$\begin{matrix}CH_3 & & CH_2\,.\,CH_2 & & H\\ & C & & C:C & \\ H & & CH_2\,.\,CH_2 & & CO_2H\end{matrix}$,

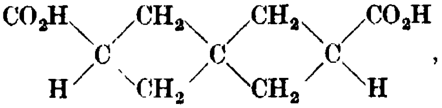
resembles the allene type II, one of the double bonds being replaced by the ring which is planar and at right angles to the plane of the adjacent double bond. This compound they synthesized,* and by resolving it into its optically active components, Perkin, Pope and Wallach† provided the first verification of van't Hoff's theory and firmly established the broad principle of the asymmetric molecule, a marked advance in stereochemistry.

---

* *Cf.* Lapworth and Wechsler, 'Trans. Chem. Soc.,' vol. 97, p. 39 (1910).

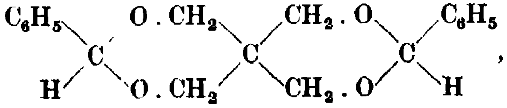
† 'Proc. Chem. Soc.,' vol. 22, p. 107 (1906).

From this work it followed that the second double bond could be replaced by a second ring and a number of investigations of dissymmetric spiran compounds have been made. The simplest of such compounds is *cyclo*butane*spirocyclo*butane-1 : 1'-dicarboxylic acid prepared by Fecht,‡



corresponding to the allene type III. This compound has been resolved by Backer and Schurink§ but it must be pointed out that the rotatory powers recorded, and those for the *dextro* component only (for the ammonium salt in water, $[\alpha]_D + 0{\cdot}13°$, and for the free acid in ether, $[M]_D + 1{\cdot}9°$), are extremely small. On the other hand, Pope and Janson‖ have recorded the complete resolution of the corresponding diamine, 1 : 1'-diamino*cyclo*butane*spirocyclo*butane, the dihydrochlorides of the optically active components of which have $[M]_{4358} \pm 30°$ in aqueous solution, and these authors have thus definitely established the dissymmetry of molecules of this type.

It has long been known¶ that compounds of this type which are most easily prepared are obtained by condensation of pentaerythritol with aldehydes (the condensation with ketones of the type *ab* : CO goes much less easily) and Read,** prepared a number of such compounds with the object of obtaining at least one of them which might be resolved into optically active components by mere crystallization. This was not realized, but Backer and Schurink (*loc. cit.*), have recorded that the dibenzylidene compound,



---

* *Cf.* Wallach, 'Liebigs. Ann.,' vol. 365, p. 266 (1909).

† 'Trans. Chem. Soc.,' vol. 95, p. 1789 (1909).

‡ 'Ber. deuts. chem. Ges.,' vol. 40, p. 3883 (1907).

§ 'Proc. K. Akad. Wetensch. Amst.,' vol. 31, p. 370 (1928).

‖ 'J. Soc. Chem. Ind.,' vol. 51, p. 316 (1932).

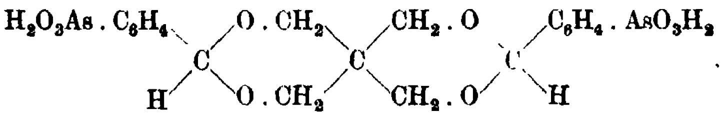
¶ Schulz and Tollens, 'Liebig's Ann.,' vol. 289, p. 28 (1896); Apel and Tollens, *ibid.*, p. 36.

** 'Trans. Chem. Soc.,' vol. 101, p. 2090 (1912).

described by Read is obtained in both *dextro-* and *laevo*-rotatory crystals having $\alpha_D \pm 2°$/mm., whereas no trace of optical activity could be detected in solution. It is now recognized from physical evidence that in such pentaerythritol derivatives the groups attached to the spiran carbon atom have a tetrahedral configuration, and, in view of the dissymmetry of the molecular structure of the particular compound, optical activity might have been expected in solution also.

A number of analogous compounds from pentaerythritol have been prepared by Böeseken and Felix.* In one case, the product obtained by hydrolysis of the condensation product of ethyl pyruvate with pentaerythritol, viz., *dl*-4 : 4′-dimethyl*spirobis*-3 : 5-dioxan-4 : 4′-dicarboxylic acid was resolved into its optically active components, having $[\alpha]_D \pm 3\cdot8°$ (acetone) and $[\alpha]_D \pm 6\cdot9°$ (water) and the optically active acids racemized rapidly in boiling aqueous solution. In another case, *dl*-4 : 4′-diaminodimethyl-*spiro*-*bis*-3 : 5-dioxan was resolved and the optical antipodes had the very small values, $[\alpha]_D \pm 0\cdot7°$ (water). These results would indicate that when such dissymmetric compounds are resolved very small rotatory powers and easy racemization are characteristic of the optically active forms.

Benzaldehyde-*p*-arsonic acid and some of its derivatives have recently been studied in some detail.† This acid undergoes condensation with pentaerythritol giving the crystalline tetrabasic *dl*-*spirobis*-3 : 5-dioxan-4 : 4′-di (phenyl-*p*-arsonic acid),



$$\begin{array}{ccccccc} H_2O_3As\,.\,C_6H_4 & & O\,.\,CH_2 & & CH_2\,.\,O & & C_6H_4\,.\,AsO_3H_2 \\ & C & & C & & C & \\ H & & O\,.\,CH_2 & & CH_2\,.\,O & & H \end{array}$$

which also belongs to the above type III. This externally compensated compound cannot be recrystallized conveniently from any solvent, but may be purified through its *salt* with two molecular proportions of *dl*-α-phenylethylamine. Although it cannot be purified from its solution in ethyl alcohol-hydrochloric acid, its solubility in this medium points to compound formation with hydrochloric acid as has been observed with other arsonic acids.‡

Combination of the above acid with two molecular proportions of *nor*-*d*-ψ-ephedrine results in the formation of two diastereoisomeric salts of which the *lAdB* is the less soluble in water and water-ethyl alcohol-acetone. Repeated recrystallization from the latter mixture is necessary to obtain the salt pure. By a similar procedure and using *nor-l-*ψ-ephedrine the pure *dAlB* salt is obtained. From these pure enantiomorphous salts, having in aqueous solution $[\alpha]_{5461}^{20^\circ} \pm 24 \cdot 2^\circ$, the two enantiomorphous acids which have, as their tetrasodium salts in aqueous solution $[\alpha]_{5461}^{20^\circ} \pm 70 \cdot 2^\circ$, were obtained. Salt formation between the arsonic acid group and suitable optically active bases has proved useful for the resolution of other externally compensated arsonic acids.* For the resolution of *dl*-N-phenylalanineamide-4-arsonic acid, Gibson and Levin† also employed *nor-d-*ψ-ephedrine

---

* 'Ber. deuts. chem. Ges.,' vol. 61, p. 787 (1928); vol. 62, p. 1310 (1929).

† Gibson and Levin, 'J. Chem. Soc.,' p. 2388 (1931); *ibid.*, p. 352 (1933).

‡ Burton and Gibson, 'J. Chem. Soc.,' vol. 125, p. 2275 (1924).

From the present work it is clear that, as Pope and Janson have shown in the case of the 1 : 1′-diamino*cyclo*butane*spirocyclo*butanes, small rotatory powers are not characteristic of the optically active isomers of this type of dissymmetric compound. The optically active acids now described in the form of their sodium salts can be boiled in aqueous solution containing a large excess of sodium hydroxide without any detectable racemization occurring and, unlike optically active compounds of this class previously described, they possess considerable optical stability.

The conversion of arsonic acids in the presence of ethyl alcohol, hydrochloric acid and a little iodine into the corresponding dichloroarsines by means of sulphur dioxide generally proceeds readily, and in the case of either of the above optically active di(arsonic acids) takes place even when the mixture is kept at 0°. The *di*(*dichloroarsine*) so obtained is a well defined substance and, in view of the stability of the optically active arsonic acids, and, bearing in mind the positions in the molecule of the arsonic acid groups, it is remarkable that its formation from the optically active acids is accompanied by complete racemization and when reconverted at the ordinary temperature into the di(arsonic acid) the latter is also optically inactive. Whatever the intermediate compound may be, it has generally been assumed that the reactions referred to concern the arsonic acid groups alone and are concisely represented

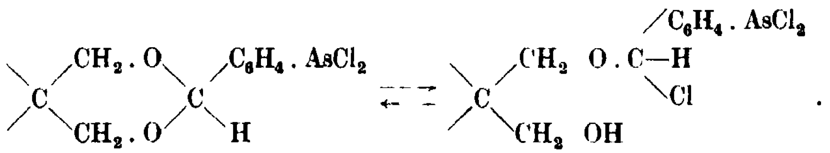
$$-AsO_3H_2 \xrightarrow[HCl]{\text{reduction}} -AsCl_2 \xrightarrow{\text{oxidation}} -AsO_3H_2.$$

The racemization which accompanies the conversion of the optically active di(arsonic acid) into the di(dichloroarsine) indicates that a more profound chemical change than the above takes place during the reduction in acid solution. The simplest explanation of the racemization would appear to be that the dioxan rings open during the reaction and then close again. The opening of the dioxan rings would destroy the dissymmetry of the molecule and hence any optical activity would disappear This explanation assumes the following change to take place if the arsonic acid groups are first reduced :

---

* Fourneau and Nicolitch, 'Bull. Soc. chim. Paris,' vol. 43, p. 1232 (1928); Gibson, Johnson and Levin, 'J. Chem. Soc.,' p. 479 (1929).

† 'J. Chem. Soc.,' p. 2754 (1929).



It may be possible to investigate these changes in cases where the intermediate compounds can be isolated.

Experimental

dl-spirobis-3 : 5-*dioxan*-4 : 4′-*di*(*phenyl*-p-*arsonic acid*) —Benzaldehyde-*p*-arsonic acid (19 gm) rapidly dissolved in a solution of sulphuric acid (30%, 53 c.c.) containing pentaerythritol (5·7 gm.) at the water bath temperature. After a colourless crystalline material had separated the mixture was heated for 30 minutes. The filtered solution was allowed to cool and the almost colourless crystalline material separated. The latter was suspended in water and treated with the calculated quantity of a standard aqueous solution of sodium hydroxide This solution was boiled up with decolorizing charcoal and from the filtrate the acid was precipitated in small colourless crystals on the slow addition of hydrochloric acid equivalent to the sodium hydroxide previously employed. While, generally, sufficiently pure for the resolution experiments, analysis showed that the acid so obtained (16 gm., 33%) was not chemically pure. Purification by direct crystallization being impracticable, it was conveniently purified through its *salt* with *dl*-α-phenylethylamine

To a suspension of *dl*-α-phenylethylamine (0·75 gm.) in water (10 c.c.) heated on the water bath, the almost pure acid (1·74 gm.) was added, complete solution taking place. To this solution ethyl alcohol (10 c.c.) and acetone (20 c.c.) was added and, after cooling, the colourless crystalline salt was separated in almost quantitative yield. It was recrystallized from a mixture of water (10 c.c.), ethyl alcohol (10 c.c.) and acetone (20 c.c.) and obtained in silky colourless needles which decomposed at 204° after previous softening (Found* :

---

* All the analyses were carried out under the direction of Dr. W. Doran in the Micro-analytical Laboratory of the University of Liverpool We gratefully acknowledge the care exercised in this work.

C 52·6, H 5·5, N 3·75, As 18·9. $C_{35}H_{44}O_{10}N_2As_2$ requires C 52·35, H 5·5, N 3·5, As 18·7%).

Pure dl-spirobis-3 : 5-*dioxan*-4 : 4′-*di* (*phenyl*-p-*arsonic acid*) was obtained by suspending the above pure *dl*-α-phenylethylamine salt in water and adding the calculated quantity of a standard aqueous solution of sodium hydroxide. This solution was extracted completely with chloroform to separate the free base and the colourless aqueous solution heated on the water bath was treated very slowly with hydrochloric acid equivalent to the sodium hydroxide previously used. The *dl*-acid separated from the hot solution in small colourless glistening crystals (undecomposed at 290°), practically insoluble in water and the usual organic solvents. The compound is, however, soluble in ethyl alcohol-hydrochloric acid but cannot be recrystallized from this medium (Found: C 40·4, H 3·72, As 27·3, $C_{19}H_{22}O_{10}As_2$ requires C 40·7, H 3·96, As 26·8%). Using King's method* the acid is tetrabasic (Found: Equiv. 142·9, 144·0. $C_{19}H_{22}O_{10}As_2$ requires Equiv. 140·0).

Resolution of dl-spirobis-3 : 5-*dioxan*-4 : 4′-di (*phenyl*-p-*arsonic acid*).—Preliminary experiments showed that *nor-d*-ψ-ephedrine (two molecular equivalents) forms a mixture of salts with the externally compensated acid (one molecular equivalent) of which the *lAdB* has a somewhat smaller solubility than its disastereoisomeride. The composition of the optically impure less soluble salt was proved by analysis (Found: C 51·2, H 5·1, N 2·95, As 17·6. $C_{37}H_{48}O_{12}N_2As_2$ requires C 51·5, H 5·6, N 3·25, As 17·4%). From these experiments the acid (10·2 gm.) having a small *laevo*-rotatory power was isolated in the usual way. This acid was added to a boiling solution of *nor-d*-ψ-ephedrine (5·5 gm) in water (250 c.c.) and this solution was evaporated on the water bath in a current of air until copious crystallization began. The colourless crystalline salt was separated from the cold solution and on recrystallization from a hot mixture of water (150 c.c.), ethyl alcohol (300 c.c.) and acetone (300 c.c.) it was obtained in small colourless needles having $[\alpha] = -4·9°$ (*c* 0·5126, water) and after five more recrystallizations from similar water-alcohol-acetone mixtures pure nor-d-ψ-*ephedrine* l-spirobis-3 : 5-*dioxan*-4 : 4′-*di* (*phenyl*-p-*arsonate*), decomposing at 225°, was obtained. It has $[\alpha] - 24·67°$† (*c* 0·4873, water) (Found:

---

* 'J. Chem. Soc.,' p. 2138 (1930). In this method aqueous solutions of barium hydroxide and barium chloride can be substituted for those of sodium hydroxide and sodium chloride respectively keeping the other conditions the same (Found: Equiv. 141·6).

† All rotatory powers were determined in 4 dm. tubes at 20° using the mercury-green line (λ 5461).

C 51·4, H 5·7, As 17·95. $C_{37}H_{48}O_{12}N_2As_2$ requires C 51·5, H 5·6, As 17·4%).*

l-spirobis-3 : 5-*dioxan*-4 : 4′-*di* (*phenyl*-p-*arsonic acid*) was obtained by dissolving the above pure *l*A*d*B salt in water, adding the calculated quantity of a standard solution of sodium hydroxide and extracting the solution with chloroform until free from *nor*-*d*-ψ-ephedrine. The aqueous solution kept at 80° was mixed slowly with the calculated quantity of an aqueous solution of hydrochloric acid when the acid separated in fine colourless needles (undecomposed at 300°). It was washed with water and dried at 110° (Found: C 40·7, H 4·45, As 27·1. $C_{19}H_{22}O_{10}As_2$ requires C 40·7, H 3·96, As 26·8%). It had [α] — 70·6° (*c* 0·4000, water containing the exact amount of sodium hydroxide to form the tetrasodium salt).

d-spirobis-3 : 5-*dioxan*-4 : 4′-*di* (*phenyl*-p-*arsonic acid*) was obtained starting with the *dl*-acid (14·4 gm.) using *nor*-*l*-ψ-ephedrine (7·75 gm.) and water (400 c.c.) following the same procedure as in the above resolution. The material separating from the aqueous solution had, after one recrystallization from the water-alcohol-acetone, mixture, [α] + 2·05° (*c* 0·5851, water). After five further recrystallizations under the same conditions it was optically pure having [α] + 23·69° (*c* 0·4908, water). *nor*-*l*-ψ-*ephedrine* d-spirobis-3 : 5-*dioxan*-4 : 4′-*di*(*phenyl*-p-*arsonate*) resembled its enantiomorph in all respects (decomposing at 225°) (Found: C 51·5, H 5·61, N 3·25, As 17·4. $C_{37}H_{48}O_{12}N_2As_2$ requires C 51·2, H 5·47, N 3·70, As 17·7%).

From this salt the pure *d*-acid was prepared in a similar manner to that used for the isolation of the *l*-acid and obtained in quantitative yield. It crystallized in fine colourless needles. Like its enantiomorph it is remarkably stable being, like it, apparently undecomposed at 300° although slight sintering had taken place. It had [α] + 69·74° (*c* 0·4666, water containing the exact quantity of sodium hydroxide to form the tetrasodium salt) (Found; C 40·8, H 4·34, As 27·1. $C_{19}H_{22}O_{10}As_2$ requires C 40·7, H 3·96, As 26·8%). From the aqueous filtrate after separating the first fraction of the above *d*A*l*B salt was obtained an optically impure *l*-acid having [α] — 31·2° under the usual conditions.

Typical racemization experiments, using the pure *l*-acid, are as follows. The *l*-acid (0·5554 gm.) was made up to 100·0 c.c. with 0·9696 N aqueous sodium hydroxide, *i.e.*, rather more than twenty-four times the amount of sodium hydroxide required to form the tetrasodium salt. This solution had [α] — 70·52° calculated on the acid content. 40·0 c.c. of this solution was boiled gently for 10 minutes during which time considerable evaporation had taken place. The solution was cooled, transferred completely with all precautions to a graduated flask and made up to 50·0 c.c. This solution had [α] — 70·34° calculated on the acid content and, clearly, no appreciable racemization had taken place.

---

* The mother liquors from this experiment should obviously have contained an excess of the *d*A*d*B salt. The *d*-acid was not obtained from these owing to an unavoidable accident ('J. Soc. Chem. Ind.,' vol. 51, p. 1064 (1932)).

The *l*-acid (1 gm.) was dissolved in ethyl alcohol (10 c.c.) and hydrochloric acid (10 c.c.) and a trace of iodine added. This solution was kept in ice-water while it was reduced with sulphur dioxide and spirobis-3 : 5-*dioxan*-4 : 4′-*di* (*phenyldichloroarsine*) separated first as an oil which quickly crystallized. After separation and drying at the ordinary temperature it was recrystallized from a mixture of benzene and ligroin (boiling point 60°–80°) and obtained in colourless crystals (melting point 163°). It was optically inactive in benzene solution (Found: Cl 23·8. $C_{19}H_{18}O_4Cl_4As_2$ required Cl 23·5%).

The di(dichloroarsine) was dissolved in benzene (15 c.c.) and shaken thoroughly at the ordinary temperature with a slight excess of a standard aqueous solution of sodium hydroxide. To the mixture an excess of an aqueous solution of hydrogen peroxide (3%) was added and, after thorough mixing, it was allowed to stand at the ordinary temperature for some 2 hours. The aqueous solution was separated, filtered and carefully acidified at the water bath temperature with hydrochloric acid. The di(arsonic acid) separated in the characteristic fine needles in theoretical quantity indicating that the oxidation was complete. The di(arsonic acid) dissolved in the calculated quantity of a standard aqueous solution of sodium hydroxide was optically inactive.

The expenses of this investigation have been met by grants which are gratefully acknowledged from the Government Grant Committee and Imperial Chemical Industries Limited.

Summary.

The molecularly dissymmetric dl-spirobis-3 : 5-*dioxan*-4 : 4′-*di*(*phenyl*-p-*arsonic acid*) has been resolved into its optically active components which have considerable rotatory powers. The optically active acids, as their sodium salts, do not undergo racemization in a boiling aqueous solution containing a large excess of sodium hydroxide. On the other hand, they undergo complete racemization when converted into the corresponding di(dichloroarsine).

Energy Relations in the β-Ray Type of Radioactive Disintegration.

By C. D. Ellis, F.R.S., and N. F. Mott.

(Received May 25, 1933.)

The difficulties connected with the continuous β-disintegration are well known. The fact that a given isotope of any element has a definite atomic weight suggests that the energy of the normal state of any nucleus is quantized; further evidence is afforded by the alternating intensities in band spectra. Evidence from the fine structure of α-rays and from the γ-rays, proves that nuclei are capable of existing in quantized excited states. In fact, in all transformations where α-particles, γ-radiation or protons are ejected from nuclei, the evidence suggests, (i) that the nuclear energy is quantized, and (ii) that energy is conserved. On the other hand, when a nucleus P transforms itself into a nucleus Q by emission of a β-particle, the β-particle has all energies between zero and a definite upper limit. One may either conclude that the energy either of P or of Q is not quantized, or that energy is not conserved in the transition. Since the α-transitions leading up to P, and starting from Q, show no sign of any indefiniteness in the energy, it is difficult to accept the former alternative; and it is thus usual to suppose that energy is not conserved.

In this paper we make the suggestion that the sharp upper limit of the β-ray spectrum is a significant parameter with which to classify a β-disintegration. We suggest the following hypotheses: two elements P, Q, such that P → Q is a β-disintegration, both possess definite atomic weights, and hence definite binding energies. Following Heisenberg* we assume that β-disintegration can only take place if the energy E_P of the nucleus P is *higher* than the energy E_Q of the nucleus Q. We make the new assumption that the energy difference $E_P - E_Q$ is *equal to the upper limit of the β-ray spectrum, i.e.*, to the maximum energy with which a β-particle can be expelled. According to our assumption, the β-particle may be expelled with *less* energy than the difference of the energies $E_P - E_Q$, of the two nuclei, but not with *more* energy. We do not wish in this paper to dwell on what happens to the excess energy in those disintegrations in which the electron is emitted with less than the maximum energy. We may, however, point out that if the energy merely disappears, implying a breakdown of the principle of energy conservation, then in a β-ray decay energy is not even statistically conserved. Our hypothesis is, of course, also consistent with the suggestion of Pauli that the excess energy is carried off by particles of great penetrating power such as neutrons of electronic mass.

---

* 'Z. Physik,' vol. 70, p. 1 (1932)

We must first modify our hypothesis to take account of the γ-radiation which follows the expulsion of the β-particle. Let us suppose that the nucleus P has energy E_P in its ground state, and that the nucleus Q formed by the expulsion of one electron has energy $E_Q(o)$ in its ground state and $E_Q(n)$ in any excited state n. (Fig. 1.) In a given fraction $p(n)$ of the disintegrations of the element P, the nucleus Q is left in an excited state n; we suggest as a further hypothesis that in these $p(n)$ disintegrations the maximum energy of the ejected particle is $E_P - E_Q(n)$. It is clear that this hypothesis provides an immediate explanation of the fact that the energies $h\nu$ of the γ-rays of a β-ray element are in general *less* than the upper limit of the β-ray spectrum. this upper limit is represented by the energy $E_P - E_Q(o)$; and no state of the element Q with energy greater than this can be excited.

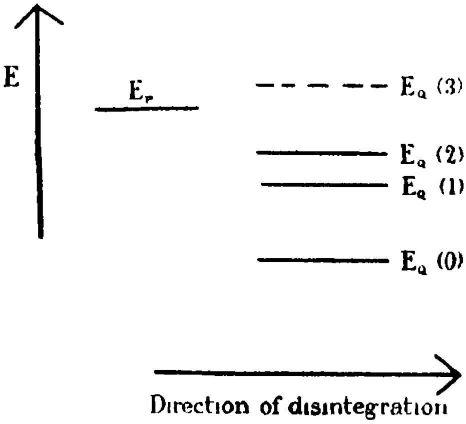


FIG. 1.

The strong γ-ray $2 \cdot 62 \times 10^6$ volts of Th C″. Pb forms an exception, being above the upper limit* $1 \cdot 8 \times 10^6$ volts of the β-ray spectrum. In this case, it is highly significant that $p(n)$ for this state is nearly unity†: in other words, an excited state of at least this energy in Th Pb is excited in nearly every disintegration. Ellis and Mott (*loc. cit.*) have also shown that it is probable that the γ-ray of energy $0 \cdot 582 \times 10^6$ volts has intensity approximately unity, and therefore that an excited state of energy $0 \cdot 582 + 2 \cdot 62 \simeq 3 \cdot 2$ million volts in Th Pb is excited at nearly every disintegration.

---

* Sargent, 'Proc. Roy. Soc.,' A, vol. 139, p. 659 (1933).

† Ellis and Mott, 'Proc. Roy. Soc.,' A, vol. 139, p. 369 (1933).

In view of these facts, we may conclude that the energy difference $E_P - E_Q(o)$ between Th C″ and Th Pb, in their normal states, is not $1 \cdot 8 \times 10^6$ volts, but $1 \cdot 8 \times 10^6 + 3 \cdot 2 \times 10^6 = 5 \cdot 0 \times 10^6$ volts. If the *direct* transition from Th C″ to Th Pb in its normal state ever takes place, the β-particles with energy up to $5 \cdot 0 \times 10^6$ are too few to have been detected with certainty.

A similar example is provided by Th B; the upper limit of the β-ray spectrum is here* $0 \cdot 36 \times 10^6$ volts. Although none of the γ-rays have energy higher than this, Ellis† has suggested that the Th C nucleus has levels at 0·238 and 0·414 million volts. On the other hand the γ-ray corresponding to the transition from $0 \cdot 238 \times 10^6$ volts to the ground state has intensity near unity.† Thus the ThC nucleus is nearly always formed in this or in higher states, and the difference in binding energy between Th B and Th C is therefore $0 \cdot 36 + 0 \cdot 238 \simeq 0 \cdot 6$ million volts.

From a consideration of the relevant data, it appears that the highest state excited in a product nucleus is usually considerably below the difference of binding energy of the β-ray nucleus and the product nucleus This is perhaps not surprising in view of the fact pointed out by Sargent (*loc. cit*) that a relation exists between the upper limit of a β-ray spectrum and the decay constant. Thus β-transitions in which the change of binding energy between the initial and final nuclei is small occur with small probability. One might expect this to occur also when the product nucleus is in an excited state, *i.e.*, if $E_P(o) - E_Q(n)$ is small, the probability of the transition is small.

We now use the value 5×10^6 volts of the energy difference between Th C″ and Th Pb deduced above to obtain a piece of evidence in favour of our assumption about the significance of the upper limit. The element Th C can become Th Pb by two branches, either

$$\text{Th C} - \beta - \text{Th C}' - \alpha - \text{Th Pb}$$

or

$$\text{Th C} - \alpha - \text{Th C}'' - \beta - \text{Th Pb},$$

and it may be assumed that the Th Pb nuclei resulting from the two processes have the same mass. If now our hypotheses about the connection between the upper limit of the β-ray spectrum and the binding energy is correct, we should expect that the difference in the binding energy between Th C and Th Pb should be equal to the *maximum* energy emitted while Th C transforms itself to Th Pb ; therefore that the *maximum* energies emitted along the two branches should be equal to one another.

---

* Sargent, *loc. cit.*

† Ellis and Mott, *loc. cit.*

In the majority of β-disintegrations of Th C, the Th C′ nucleus is left unexcited, and we may safely conclude, according to our hypothesis, that the energy, $2 \cdot 2 \times 10^6$ volts, of the observed upper limit of the β-ray spectrum is equal to the difference in the binding energy of the two unexcited nuclei The energy of the α-particle from Th C′ (corrected for the recoil) is $8 \cdot 95 \times 10^6$, and of the fastest α-particle from Th C $6 \cdot 20 \times 10^6$ volts. It is now well established that both these values represent the total energy of disintegration. Thus the maximum energy given out along the Th C . C′ . Pb branch is $2 \cdot 2 + 8 \cdot 95 = 11 \cdot 15$ volts $\times 10^6$ and along the branch Th C . C″ . Pb $6 \cdot 20 + 5 \cdot 0 = 11 \cdot 20$ volts $\times 10^6$. The agreement is very satisfactory, but we do not wish to over-emphasize the support which it yields to our hypothesis, since the data about the upper energy limits of the β-ray bodies are not yet of a high order of accuracy.

Note that, according to our hypothesis, a long range α-particle is only emitted following a β-disintegration of considerably less than the maximum energy The excess energy of a long range α-particle cannot be greater than the upper limit of the β-ray spectrum

§ 2. *The Shape of the Continuous β-ray Spectrum.*

We have already introduced the quantity $p(n)$, which is equal to the probability that a given nucleus is formed in an excited state n, after a β-disintegration. $p(n)$ is of the dimensions of a number, and

$$\Sigma_n p(n) = 1$$

Let the probability that a β-particle is ejected from the parent element with energy between E and E + dE, leaving the product nucleus in the state n, be some function $f_n(\mathrm{E})\,d\mathrm{E}$, so that

$$\int f_n(\mathrm{E})\,d\mathrm{E} = p(n)$$

Then the observed number of β-particles with energies between E E + dE, will be per disintegration

$$\Sigma_n f_n(\mathrm{E})d\mathrm{E}.$$

Now according to our hypothesis, f_n (E) vanishes if E is greater than some maximum value E_n, given by, fig. 1,

$$\mathrm{E}_n = \mathrm{E}_{\mathrm{P}} - \mathrm{E}_{\mathrm{Q}}(n).$$

Let us make the further hypothesis that the form of the curve corresponding to a definite mode of disintegration is always the same, thus we write

$$f_n(E) = Cf(E/E_n),$$

where f is the same for all elements and all n, and C is adjusted so that

$$C\int_0^{E_n} f(E/E_n)\,dE = p(n).$$

This hypothesis is not, of course, based on any theory, since we have no idea of the cause of the continuous spectrum; we adopt it provisionally because of its simplicity

Since no γ-rays are emitted after the disintegration of Ra E, we assume that the Po nucleus is always formed in its ground state. Thus the continuous β-ray spectrum for Ra E may be taken to be that typical of a single mode of disintegration, and we have based the calculations which follow on the curve given recently by Sargent.*

It is easily seen that to obtain the contribution to the nth mode (maximum energy E_n) of any other disintegration, we must multiply the abscissæ of the Ra E curve by E_n/E_0, and the ordinate by $p(n)E_0/E_n$.

We now apply these ideas to the Ra C . C′ disintegration, where we have considerable information about the excited states $E_0(n)$ in which the Ra C′ nucleus is left, and also about the intensities $p(n)$. The most direct evidence about the energies of the excited states is provided by the groups of the long range α-particles; we are grateful to Lord Rutherford for informing us that in a recent investigation in collaboration with Wynn-Williams, Lewis and Bowden it has been possible to detect several new long-range groups in addition to those found in earlier work.

For our purpose we need only consider those excited states which are excited often, so that $p(n)$ is comparable with unity. Now $p(n)$ is equal to the sum of the numbers of long range α-particles, γ-quanta and associated β-particles per disintegration starting from the state n; and in general the number of γ-quanta is much the greatest of these. It follows that we need only consider those excited states which are associated with strong γ-ray emission.

The more prominent γ-rays are shown in Table I, together with their intensities in quanta per disintegration.†

---

* 'Proc. Camb. Phil. Soc.,' vol. 28, p. 550 (1932).

† Ellis and Aston, 'Proc. Roy. Soc.,' A, vol. 129, p. 180 (1930). The energies are based on some new measurements which will be published shortly.

We are here again indebted to Lord Rutherford, Wynn-Williams, Lewis and Bowden for informing us that these γ-rays fit in excellently with their new measurements of the long range α-groups and they have given us their provisional values for the energies of some of the groups. From a consideration of these data and of both the energies and intensities of the γ-rays it appears probable that the excitation scheme of the Ra C′ nucleus is as shown in Table II. Precise values of the excitation energies are not necessary for our present purpose and we have therefore given rounded values to avoid confusion with the final results when they are published by the above authors. It is also not necessary for us to discuss in detail how the γ-rays fit into this scheme, except to mention that γQ starts like γW from the 2·70 level and finishes on the level which is the end of γR and the beginning of γV.

Table I.—Ra C . C′ γ-rays.

Name.	$h\nu$ volts × 10^{-6}.	p Quanta per disintegration.
γNa	0·426	*
γNb	0·498	*
γO	0·607	0·66 (approx.)
γPa	0·766	0·065
γQ	0·933	0·065
γR	1·120	0·21
γS	1·238	0·065
γSa	1·379	0·065
γV	1·761	0·26
γW	2·198	0·074

* Not measured but suspected of order 0·03 to 0·06.

Table II.

Excitation energy E(n) of level, volts × 10^{-6}.	γ-rays given by transition to ground.	Excitation probability $p(n)$.
0·61	γO	0·66
1·67	γNa + γS	0·065
2·14	γPa + γSa	0·065
2·70	γNb + γW	0·14
2·88	γR + γV	0·21

The sum of the excitation probabilities is 1·1, suggesting that the measured intensities are rather high, but we need not discuss this point, since for the present purpose we only require relative intensities. However, it seems rather unlikely that the Ra C′ nucleus is often formed in the unexcited state and we shall assume therefore that the observed upper limit of the β-ray spectrum ($3 \cdot 15 \times 10^6$ volts) corresponds to the formation of a Ra C nucleus in the excited state of $0 \cdot 61 \times 10^6$ volts excess energy. Thus the total energy of disintegration (difference of binding energy of Ra C and Ra C′) is $E_0 = 3 \cdot 15 + 0 \cdot 61 \simeq 3 \cdot 76$ million volts. To obtain the complete spectrum we have therefore to add together five continuous spectra similar to that of Ra E, with end points equal to $(E_0 - E(n))$, where $E(n)$ is the excitation energy, and weighted in proportion to the intensity $p(n)$ of excitation

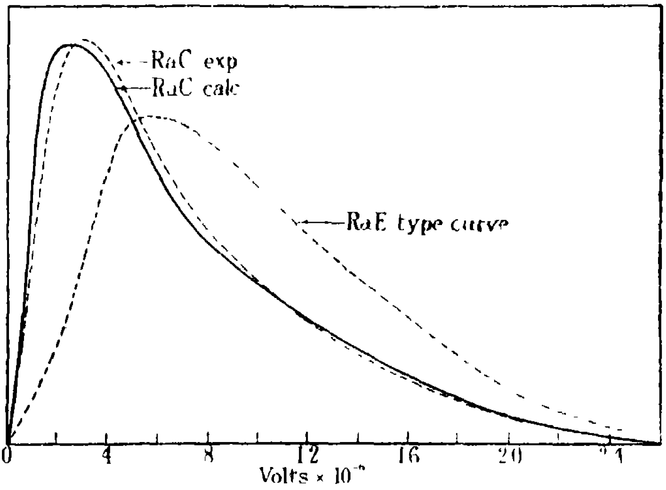


FIG. 2.

The result is shown in fig 2, it can be seen that the calculated curve is very similar to the experimental curve, which is due to Gurney.* It must be emphasized that the upper limit of the calculated curve is of course adjusted to fit the experimental, and that the two curves are drawn to have the same area. The test of the theory lies in its ability to predict the position of the maximum. To give a basis for estimating the agreement, a simple curve of Ra E type with an upper limit of $3 \cdot 15 \times 10^6$ volts and the same area is also shown.

The excitation has here only been treated in an approximate way, since there are a large number of weak γ-ray lines which we have not considered. However, more detailed investigations would hardly be justified at present since our data about the Ra E and Ra C β-ray curves is uncertain. The theory emphasizes the importance of fresh measurements of these continuous spectra.

---

* 'Proc. Roy. Soc.,' A, vol. 109, p. 540 (1925).

Similar considerations may be applied to Th B . C. We have shown (*loc. cit.*) that the level in Th C of $0 \cdot 238 \times 10^6$ volts is excited* 0·96 times per disintegration, all other excitation probabilities being less than 0·05. The upper limit of the β-ray spectrum, according to Sargent (*loc. cit.*) is $0 \cdot 36 \times 10^6$ volts. We deduce that the difference of energy of Th B and Th C is $0 \cdot 36 + 0 \cdot 238 \simeq 0 \cdot 6$ million volts, as stated in § 1 , if any β-rays of this energy are present corresponding to a direct transition to the ground state, they are obscured by the Th C spectrum. The continuous β-ray spectrum should be similar to that of Ra E, in view of the predominance of one mode of disintegration. The data available are not sufficiently accurate to justify a detailed comparison, but it is interesting to note that the relative positions of the maximum and end point of the Th B spectrum lend support to this view. The data is given in Table III.

Table III.

	Position of maximum volts × 10^{-5}	End point volts × 10^{-5}.	Ratio end point/maximum.
Ra E	$2 \cdot 2 \times 10^5$	$12 \cdot 5 \times 10^5$	0·18
Th B	$0 \cdot 66 \times 10^5$	$3 \cdot 6 \times 10^5$	0·18

The two bodies Th C and Th C″ are interesting because for both we have a fairly accurate knowledge of the states of excitation in which the product nucleus is left. Th C″ . Pb is the simplest disintegration because it seems probable that in nearly every disintegration the lead nucleus is left with $3 \cdot 2 \times 10^6$ excitation energy.† To a close approximation we should therefore expect the continuous spectrum to be similar to that of Ra E, the upper limit being at $1 \cdot 8 \times 10^6$ volts and the ordinates adjusted correspondingly. As already mentioned, the total energy of disintegration E_0 will be $(1 \cdot 8 + 3 \cdot 2) \times 10^6$ volts $= 5 \cdot 0 \times 10^6$ volts, and, according to our hypothesis, there may be an additional faint continuous spectrum with an upper limit of this value, corresponding to the rare case of the lead nucleus being left unexcited. It is possible that this is the explanation of the weak band from Th (C + C′ + C″) detected by magnetic analysis by Yovanovitch and d'Espine* which had an upper limit of about 4·9 × 10^6 volts. There is at present no means of testing the correctness of our conclusion about the shape of the main curve since this has not been measured. It is, however, well established† that the β-rays of Th C″ . Pb show an exponential absorption in aluminium, and this is some indication that the curve is similar to that of Ra E, and that there are not an unusual number of slow electrons. We hope that experiments will be carried out to test this point very shortly.

---

* In our paper we give the value 1·22, using the value of the internal conversion coefficient calculated by Taylor and Mott. If one takes the experimental value for a Ra B line of this frequency, we obtain the value quoted. A value greater than one is, of course, impossible.

† Ellis and Mott, *loc. cit.*

In the majority of the Th C . C′ disintegrations, the Th C′ nucleus is formed unexcited, as is shown by the absence of any strong γ-ray following this disintegration, this mode of disintegration will account for the bulk of the continuous β-spectrum. In a small number of disintegrations, however, the nucleus is left excited, as shown by the existence of the long range α-particle groups We are indebted to Lord Rutherford, Lewis and Bowden, for giving us their latest data about these long range groups, and this combined with some recent measurements of the β-ray spectrum, suggests the following excitation scheme.

Table IV.

Excitation energy, volts × 10^{-6}.	Excitation probability, $p(n)$.	Corresponding upper limit of β-ray spectrum, volts × 10^{-6}.
0	0·75	2·2
0·726	0·2	1 5
1 800	0·05	0 4

We have computed the resultant continuous spectrum to be expected and combined it with a simple " Ra E " type curve for the Th C″ disintegration, approximately weighted to correspond to the branching at Th C. The result is shown in fig. 3.

This spectrum has been measured by Gurney,‡ and his experimental curve drawn to the same area is shown dotted in fig. 3 , and it will be seen that there is a marked disagreement between the two curves. We do not feel at all convinced, however, that the experimental curve is correct. It is not known with anything near the accuracy of the Ra C curve, since only relatively small radioactive sources were available for the measurement, and very wide defining slits were used. Also the predominance of slow electrons in the experimental curve seems not to be in agreement with the approximate exponential absorption of the β-rays in aluminium. We feel that a decision may fairly be deferred until new measurements have been made.

---

* 'J. de Physique,' vol. 8, p. 276 (1927).

† Marsden and Darwin, 'Proc. Roy. Soc.,' A, vol. 87, p. 17 (1912); Hahn and Meitner, 'Z Physik,' vol. 13, p. 380 (1912).

‡ 'Proc. Roy. Soc.,' A, vol. 112, p. 380 (1926).

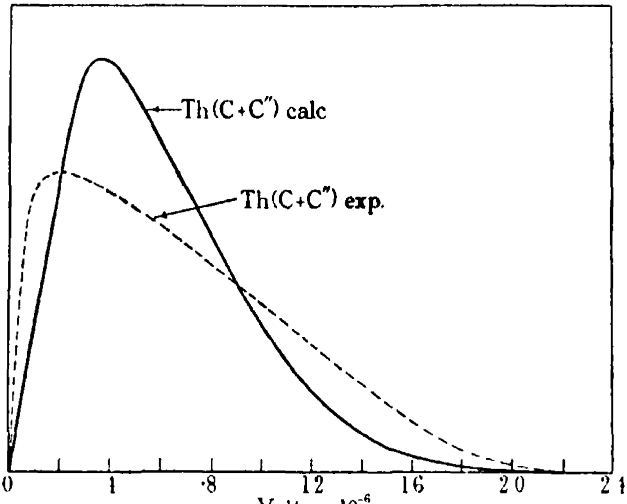


FIG. 3.

We have discussed the cases of Th C and Th C′ mainly to show how our hypothesis makes it possible to correlate data on γ-rays and the continuous spectrum curves, and in this way gives a real point and interest to further investigation of continuous β-ray spectra.

Summary.

It is suggested that the upper limit of the continuous β-ray spectrum corresponds to the difference of binding energies of the initial and final nuclei; experimental evidence is given in favour of this hypothesis. It is further suggested that the observed continuous β-ray spectra may be analysed into several curves superimposed, all of which have the same shape; this hypothesis is applied to the β-rays from Ra E, Ra C, Th B, Th C, Th C″. Some new objects for experimental investigation are suggested.

Ocean Currents Produced by Evaporation and Precipitation.

By G R. Goldsbrough, F.R.S.

(Received June 10, 1933.)

At the end of his paper on the long period tides Hough* discussed the effects of precipitation and evaporation in reference to steady currents in the ocean. He assumed that the ocean covered the whole globe and that the precipitation and evaporation were functions of the latitude only. His work showed that no steady motions exist in this case.

The present paper examines the case of an ocean bounded by shores extending from pole to pole along meridians, for which the evaporation and precipitation are also functions of the longitude. The results show that for a probable system of distribution of evaporation and precipitation steady currents do exist and are of the kind actually observed in the great oceans.

Consider the motion of a homogeneous liquid forming a spherical layer of approximately uniform depth upon a solid rotating globe. Let u be the southward component of velocity and v the eastward component of velocity at a point in co-latitude θ and longitude ϕ. If there is no external disturbing force and the motions are small, the dynamical equations are

$$\left.\begin{aligned}\frac{\partial u}{\partial t} - 2\omega v\cos\theta &= -\frac{g}{a}\frac{\partial\zeta}{\partial\theta}\\ \frac{\partial v}{\partial t} - 2\,\omega u\cos\theta &= -\frac{g}{a\sin\theta}\frac{\partial\zeta}{\partial\phi}\end{aligned}\right\} \qquad (1)$$

where ζ is the displacement of the free surface. For steady motions these reduce to

$$\left.\begin{aligned}2\omega v\cos\theta &= \frac{g}{a}\frac{\partial\zeta}{\partial\theta}\\ 2\omega u\cos\theta &= -\frac{g}{a\sin\theta}\frac{\partial\zeta}{\partial\phi}\end{aligned}\right\}, \qquad (2)$$

and ζ is now independent of the time.

For the equation of continuity we make the usual assumptions of the tidal theory, that the horizontal motion is the same at all points in a vertical line and that the vertical motion is small compared with the horizontal motion.

---

* 'Phil. Trans.,' A, vol. 189, p. 253 (1897).

Consider a column of liquid standing on a base $a \sin \theta\delta\phi \times a\delta\theta$ and of height $h + \zeta$, where h is the undisturbed depth of the free surface, assumed constant. Under steady conditions the quantity of liquid in this column must remain unchanged. The rate of flow of the liquid out of the column will be

$$\frac{\partial}{\partial\theta}(hua \sin \theta\delta\theta\delta\phi) + \frac{\partial}{\partial\phi}(hva\delta\theta\delta\phi),$$

neglecting ζ compared with h.

Now let there be a function $P(\theta, \phi)$ per unit area applied to the surface indicating that, when P is positive, liquid is abstracted from the surface and when P is negative, liquid is deposited upon the surface. Thus a positive P implies evaporation at a rate given by the function and negative P implies precipitation. The operation is also supposed to be such as not to impair the dynamical equations already stated.

The rate of flow of the liquid out of the column just mentioned will then be $Pa^2 \sin \theta\delta\theta\, \delta\phi$. Hence we have

$$\frac{\partial}{\partial\theta}(hu \sin \theta) + \frac{\partial}{\partial\phi}(hv) + Pa \sin \theta = 0. \tag{3}$$

On substituting $\mu = \cos \theta$ and solving equations (2) and (3) we find

$$u = -\mu(1 - \mu^2)^{-\frac{1}{2}}Pa/h, \tag{4}$$

$$\frac{\partial v}{\partial\phi} = -\frac{\sqrt{(1 - \mu^2)}\, a}{\mu h}\frac{\partial}{\partial\mu}(\mu^2 P), \tag{5}$$

$$\frac{\partial\zeta}{\partial\phi} = 2\omega a^2\mu^2\, P/gh. \tag{6}$$

Certain boundary conditions have also to be fulfilled. We shall consider an ocean bounded from pole to pole by two meridional coasts at $\phi = 0, \alpha$ Then we must have $v = 0$ when $\phi = 0, \alpha$, and $\sqrt{(1 - \mu^2)}u \rightarrow 0$ was $\mu \rightarrow \pm 1$.

The equation of the stream lines is

$$d\theta$$

On substituting for u, v we have

$$\frac{\mu d\mu}{\partial\zeta/\partial\phi} = -\frac{\mu d\phi}{\partial\zeta/\partial\mu}.$$

This gives $\mu = 0$ or $\zeta = \text{constant}$. The stream lines are therefore along the contours of the disturbed surface, except that the equator is always a stream line.

Certain deductions follow at once from (4), (5), (6). The velocity components are independent of the value of the angular velocity of rotation, though it is clear that they are only determinate when $\omega \neq 0$. The surface displacement ζ varies as ω.

Since we must have $v = 0$ when $\phi = 0, \alpha$, it follows from (5) that

$$\int_0^\alpha \mathrm{P}\, d\phi = 0. \tag{7}$$

Now consider the area of ocean surface bounded by two parallels of latitude μ and $\mu + \delta\mu$ and terminated by the meridians $\phi = 0, \alpha$. The result (7) means that over the whole of this area there is no gain or loss of liquid due to the function P. The necessity of fulfilling this condition places a considerable restriction upon the function P and prevents it from being entirely arbitrary. On the other hand the condition is evidently rigorous for systems of this type.

Examples of Steady Motion in Oceans under these Conditions.

As a simple illustration of motion due to the operation of the function P on the surface consider

$$\mathrm{P}a/h = k\sqrt{(1 - \mu^2)} \cos \pi\phi/\alpha. \tag{8}$$

This function satisfies the condition (7) and the factor $\sqrt{(1 - \mu^2)}$ is included so that u may not be infinite at the poles. We have then

$$u = -k\mu \cos \pi\phi/\alpha, \tag{9}$$

$$v = -k\,(a\alpha/h\pi)\,(1 - 2\mu^2) \sin \pi\phi/\alpha, \tag{10}$$

$$\zeta = k\,(2\omega a^2\alpha/\pi g h)\,\mu^2 \sqrt{(1 - \mu^2)} \sin \pi\phi/\alpha + \text{const.} \tag{11}$$

The constant in (11) is determined from the fact that the integral of ζ over the whole surface of the ocean must be zero.

Examining now the velocity components as given in (9) and (10), we see that u is negative at $\phi = 0$ and positive at $\phi = \alpha$. That is, the current moves northward on the western boundary and southward on the eastern boundary.

In a similar manner we see that when $\mu^2 < \frac{1}{2}$, v is negative across the whole ocean. But when $\mu^2 > \frac{1}{2}$, it is positive across the whole ocean. The motion is therefore a circulation of the ocean in the northern hemisphere about the point $\mu = 1/\sqrt{2}$, $\phi = \frac{1}{2}\alpha$, and in a clockwise manner as seen from above the ocean. The motion in the other hemisphere is the image about the equator of that in the first.

The form of the function P giving rise to this motion represents evaporation over the western half and precipitation over the eastern half of the ocean, the rate of each diminishing to zero at the poles.

As a second example take

$$Pa/h = k\sqrt{(1 - \mu^2)}\cos 2\pi\phi/\alpha. \tag{12}$$

From this

$$u = -k\mu\cos 2\pi\phi/\alpha, \tag{13}$$

$$v = k\,(a\alpha/2h\pi)\,(1 - 2\mu^2)\sin 2\pi\phi/\alpha \tag{14}$$

The distribution (12) consists of evaporation from the ocean between $\phi = 0$ and $\phi = \frac{1}{4}\alpha$, precipitation between $\phi = \frac{1}{4}\alpha$ and $\phi = \frac{3}{4}\alpha$, and evaporation again from $\phi = \frac{3}{4}\alpha$ to $\phi = \alpha$.

The velocity component v vanishes at the boundaries and in addition at $\phi = \frac{1}{2}\alpha$. It is evident therefore that the ocean is now divided into four parts by the equator and the central meridian, in each of which the currents are self-contained. The systems are images of one of them in the lines $\mu = 0$, $\phi = \frac{1}{2}\alpha$

Application to the North Atlantic Ocean

By making certain assumptions we may find the contribution of evaporation and precipitation to the production of steady currents in the North Atlantic Ocean. In the absence of large scale air movements it may be taken that the evaporation and precipitation of ocean water are functions of the latitude only. The imposition of a strong westerly wind on the western shore of the ocean (such as actually occurs during the winter months) will tend to reduce the precipitation and increase the evaporation along that shore.* We may therefore represent the effect by choosing the function P in such a way that it is positive in the vicinity of $\phi = 0$ and negative over the rest of the ocean always subject to the condition

$$\int_0^\alpha P\,d\phi = 0.$$

We therefore take

$$\begin{aligned} Pa/h &= k\,[(\phi - n\alpha)^2 - \tfrac{1}{3}n^3\alpha^2]\,(1 - \mu^2)^3, && 0 \leqslant \phi \leqslant n\alpha, \\ &= -\tfrac{1}{3}kn^3\alpha^2\,(1 - \mu^2)^3, && n\alpha \leqslant \phi \leqslant \alpha, \end{aligned}$$

where n is a fraction.

The latitude factor $(1 - \mu^2)^3$ is selected for reasons that will appear later.

---

* See Lake, "Physical Geography," p. 102 (1929).

By use of equations (4), (5), (6) we find

$$
\begin{aligned}
u &= -k\mu(1-\mu^2)^{5/2}\{(\phi-n\alpha)^2-\tfrac{1}{3}n^3\alpha^2\}, && 0 \leqslant \phi \leqslant n\alpha. \\
&= -\tfrac{1}{3}k\mu(1-\mu^2)^{5/2}\,n^3\alpha^2, && n\alpha \leqslant \phi \leqslant \alpha, \\
v &= -2k(1-\mu^2)^{5/2}(1-4\mu^2)\{\tfrac{1}{3}(\phi-n\alpha)^3-\tfrac{1}{3}n^3\alpha^2\phi+\tfrac{1}{3}n^3\alpha^3\}, && 0 \leqslant \phi \leqslant n\alpha \\
&= -2k(1-\mu^2)^{5/2}(1-4\mu^2)\{-\tfrac{1}{3}n^3\alpha^2\phi+\tfrac{1}{3}n^3\alpha^3\}, && n\alpha \leqslant \phi \leqslant \alpha, \\
\zeta &= (2\omega ak/g)\,\mu^2(1-\mu^2)^3\{\tfrac{1}{3}(\phi-n\alpha)^3-\tfrac{1}{3}n^3\alpha^2\phi+\tfrac{1}{3}n^3\alpha^3\}, && 0 \leqslant \phi \leqslant n\alpha \\
&= (2\omega ak/g)\,\mu^2(1-\mu^2)^3\{-\tfrac{1}{3}n^3\alpha^2\phi+\tfrac{1}{3}n^3\alpha^3\}, && n\alpha \leqslant \phi \leqslant \alpha.
\end{aligned}
\tag{15}
$$

The form of the latitude factor $(1-\mu^2)^3$ in P was selected so that u would not be infinite at the poles and so that v would vanish at $\mu = \pm\frac{1}{2}$, that is in latitude ± 30°, in accordance with the observations of currents in both the north and south Atlantic Ocean.

As before it is noticeable that the currents produced in this way are clockwise to an observer in the northern hemisphere above the earth's surface.

The fraction n is arbitrarily taken as 1/5 and the curves ζ = constant drawn The diagram shows these curves in orthographic projection.

The resemblance to the circulation of the North Atlantic Ocean will be seen at once. The currents in the diagram have the same direction as those of the ocean, the strongly marked current along the western boundary resembles the Gulf Stream and there is a definite set of the current along the equator.

Throughout it has been assumed that the depth of the ocean was a constant h and that the whole of the ocean from the surface to the bed was in motion. In making a comparison with the actual Atlantic currents we shall take h as being the depth of the moving liquid. (In any precise estimate of the speed of the current account would have to be taken of the loss of energy through frictional resistances and the problem would be more complicated.) We have then from (4), that for latitude 45° and at the shore $\phi = 0$, the current speed is $\mathrm{P}a/h$.

Hence for a current of one mile an hour, P will be $h/4000$ feet per hour, where h is the depth of the moving liquid in feet. If we take h at 100 feet (and this is probably too small even for a mean value), P would be about 7 inches per day. That is to say, to maintain a current of one mile per hour at the selected point, there would have to be a difference of 7 inches per day between the evaporation and precipitation at that point. This requirement is excessive. The investigation, however, does show that precipitation and evaporation give rise to steady currents and that these currents are of the type observed in the ocean.

Summary.

In this paper the effects of precipitation and evaporation at the surface of an ocean on a rotating globe in reference to the production of steady currents are considered. The ocean is of uniform depth and bounded by two coasts extending from pole to pole along meridians of longitude. When the precipitation and evaporation are distributed over the surface of this ocean in a manner comparable with the probable distribution over the Atlantic Ocean, it is shown that steady currents exist, having a character and direction similar to those observed in the Atlantic Ocean, though not of the same magnitude.

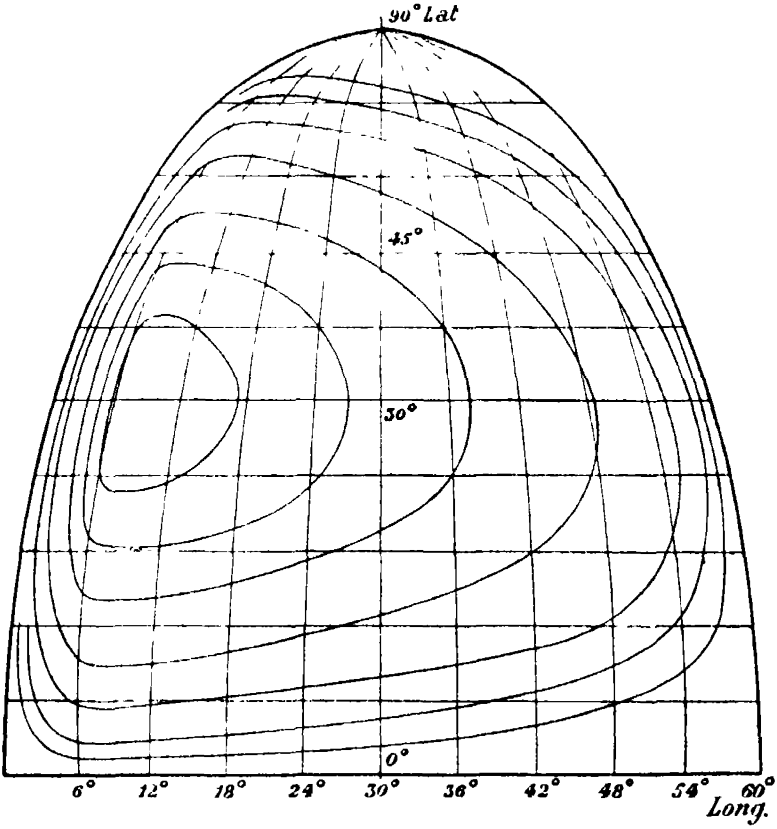


FIG. 1.

Probability and Chance in the Theory of Statistics.

By M. S. Bartlett, B.A., Queens' College, Cambridge.

(Communicated by G. Udny Yule, F R.S.—Received March 10, 1933.)

1. *The distinction between probability and chance.*—Probability is of fundamental importance in statistical theory, which indeed we may regard merely as part of the mathematical theory of probability, yet statisticians still differ widely in opinion, among themselves as well as with others, on what it means and on how it should be used. One point that is therefore worth stressing is the distinction between probability and chance. Chance is naturally included in the wider field of probability, but at times some confusion seems to have arisen through a failure to distinguish between the two This confusion has been related also to the fact that the probability of an event is not unique, but must depend on the data on which it is based.

It is of more importance here to indicate the difference in meaning between probability and chance than to define each term individually. There is much to be said for leaving probability as an undefined concept, particularly as all attempts at definition can be criticized in one way or another, and it will thus be assumed sufficient to say* that the probability P of a proposition or of an event a, on a set of data or premisses d, is the degree of rational belief in a, given d.

Some writers, Bayes† and Venn,‡ for example, have regarded the term chance as synonymous with probability. It may be remarked, however, that a chance is of a more objective character than other probabilities (no probability is of course objective in the sense that it denies causality, with the possible exception of probabilities in the quantum theory). Once given a particular premiss, it may be possible to give a probability a value with which everyone readily agrees; this value is, moreover, *independent of further data or premisses* unless these contradict the initial premiss, that is, all other knowledge is irrelevant. The probability may then be said to be a chance. This invariant character of a chance depends always on the recognition that the particular premiss essential to it is inserted in the data.

---

* *Cf.* Keynes, "A Treatise on Probability," p. 4, (1921).

† 'Phil. Trans.,' vol 53, p 370 (1763).

‡ "The Logic of Chance," 3rd ed., p. 162 (1888).

For example, the chance of an ordinary die giving 6 at a single throw, with the premiss that it is unbiassed, is $\frac{1}{6}$, and is independent of how often it has fallen 6 before; the probability merely on the data that the particular die considered has been observed to give 6 with suspicious facility in the past, would be greater than $\frac{1}{6}$. The notation for probabilities to be employed here is that used by Jeffreys.* The probability of a, given d, is written $\mathrm{P}(a \mid d)$. One modification is made, that a probability may be denoted by p if the fact that it is a chance requires emphasis. Thus

$$p(6 \mid f, d) = p(6 \mid f) = \tfrac{1}{6}$$

denotes that the chance of 6 with the premiss f that the die is unbiassed is $\frac{1}{6}$, whatever the other data d, whereas $\mathrm{P}(6 \mid d)$ denotes the probability of the die giving 6 on the data d only, and this may vary as d varies.

For a continuous variable x, it is convenient to speak of the probability of x when we mean more precisely the probability of a value of x in the range $(x, x + dx)$. Thus, if we assume the chance of x, the measure of an observation, to be given by the normal law, a premiss f, and we assume further that the true mean is m, and the true standard deviation is σ, we have

$$p(x \mid f, m, \sigma) = (2\pi\sigma^2)^{-\frac{1}{2}} e^{-\frac{1}{2}(x-m)^2/\sigma^2} dx.$$

2. *Probability and Frequency.*—There have been many attempts to define the probability of an event in terms of frequency, as the limit of the ratio of the number of successes r to the number of trials n, as n is made indefinitely large. This definition can have no meaning unless it is confined to discussions on chance, and then it is neither necessary nor satisfactory. It may appeal at first because it sounds concrete, but when we reflect (1) that we could never know the true value of a probability until we had made an infinite number of trials; (2) that actual experiments with dice have often not caused us to conclude that our estimate *a priori* of a probability of any number with a true die is wrong, but that something was wrong with the dice used, and (3) that we are always considering probabilities on hypothetical premisses, then the definition no longer appears so real. Indeed, it is obvious that since a probability is not unique, but depends on the data, the ratio of frequencies must be a ratio of hypothetical frequencies depending also on the data, and any concreteness the definition once seemed to have entirely disappears. Fisher†

---

* "Scientific Inference," p. 15 (1931).

† 'Phil. Trans.,' A, vol. 222, p. 309 (1922).

would appear to favour this kind of definition, though he recognizes its limitations, for he stipulates that the "population" of trials shall be both infinite and hypothetical.

It is worth while examining this notion of a limit, which is related to one of the first and most famous theorems in statistics, Bernoulli's theorem of large numbers. This states that if the chance of an event is p, we can, given any quantity ε, make the chance $P(n)$, say, that

$$|p - r/n| > \varepsilon$$

as small as we please by taking a sufficiently large n.

This would, of course, follow automatically if we defined the probability of the event as the limit of the ratio r/n, but as a statement about the limit of r/n on the *a priori* notion of probability, it is, as Wrinch and Jeffreys* have pointed out, hardly sufficient, for we must consider the probable value of $|p - r/n|$ not only for a particular n, but for all $n > n_0$, say.

We have the chance of r successes in n trials given by

$$f(r), \quad \text{say}, \quad = {}^nC_r p^r q^{n-r}, \tag{1}$$

where $q = 1 - p$. Then for n large, we may write†

$$f(r) = (2\pi npq)^{-\frac{1}{2}} e^{-\frac{1}{2}\alpha^2/pq} \{1 + 0(n^{-\frac{1}{2}})\}, \tag{2}$$

where $r = np + \alpha\sqrt{n} + \eta$, $|\eta| < k$.

Writing $r - np = x(nqp)^{\frac{1}{2}}$, we can conveniently replace the summation of terms $f(r)$ by integration with respect to x, and obtain that the chance that $|p - r/n| > \varepsilon$ is

$$P(n) = \lambda \int_{x_0}^{\infty} e^{-\frac{1}{2}x^2}\, dx, \tag{3}$$

where $\lambda < C$ for $n > N$, say, $C - 2(2\pi)^{-\frac{1}{2}}$ being small for N large, and $x_0 = \varepsilon(n/pq)^{\frac{1}{2}}$. Integrate by parts, then

$$P(n) < C \left\{ \left. \frac{-e^{-\frac{1}{2}x^2}}{x} \right|_{x_0}^{\infty} - \int_{x_0}^{\infty} \frac{e^{-\frac{1}{2}x^2}}{x^2}\, dx \right\},$$

$$< C(pq)^{\frac{1}{2}} e^{-\frac{1}{2}\varepsilon^2 npq}/(\varepsilon\sqrt{n}). \tag{4}$$

That $P(n) \to 0$ as $n \to \infty$ is Bernoulli's theorem. But we require to consider the chance that $|p - r/n| > \varepsilon$ for at least one $n > n_0$. We must, of course, for each n consider *all* the n trials in assessing this chance, which is $P'(n_0)$, say.

---

* 'Phil. Mag.,' vol. 38, p. 715 (1919).

† "Scientific Inference," p. 240.

Let $n_s = n_0 + s$, $\mathrm{P}(\sim n_s)$ denote the chance that '$|p - r/n| > \varepsilon$' is not true for $n = n_s$, $\mathrm{P}(n_t, \sim n_s)$ the chance that it is true for $n = n_t$, not true for $n = n_s$, and so on, then

$$\mathrm{P}'(n_0) = \mathrm{P}(n_1) + \mathrm{P}(n_2, \sim n_1) + \mathrm{P}(n_3, \sim n_2, \sim n_1) + \ldots,$$

$$\leq \mathrm{P}(n_1) + \mathrm{P}(n_2) + \mathrm{P}(n_3) + \ldots,$$

$$\leq \frac{\mathrm{C}(pq)^{\frac{1}{2}}}{\varepsilon} \sum_{n=n_1}^{\infty} \left\{ \frac{e^{-\frac{1}{2}\varepsilon^2 npq}}{\sqrt{n}} \right\}.$$

Since each term is less than its integration with respect to n in the corresponding range $(n-1, n)$, we have

$$\mathrm{P}'(n_0) < \frac{\mathrm{C}(pq)^{\frac{1}{2}}}{\varepsilon} \int_{n_0}^{\infty} \frac{e^{\ \frac{1}{2}\varepsilon^2 npq}}{\sqrt{n}}\, dn.$$

Integrate by parts, then

$$\mathrm{P}'(n_0) \leq \frac{2\mathrm{C}}{\varepsilon^3 (pq)^{\frac{1}{2}}} \left\{ \frac{-e^{\ \frac{1}{2}\varepsilon^2 npq}}{\sqrt{n}} \Bigg|_{n_0}^{\infty} - \int_{n_0}^{\infty} \frac{e^{-\frac{1}{2}\varepsilon^2 npq}}{2n}\, dn \right\}$$

$$\leq 2\mathrm{C}(pq)^{-\frac{1}{2}} e^{-\frac{1}{2}\varepsilon^2 n_0 pq}/(\varepsilon^3 \sqrt{n_0}). \qquad (5)$$

Thus though when ε small, $\mathrm{P}'(n_0)/\mathrm{P}(n_0)$ is large, we can by taking a sufficiently large n_0, make P' as well as P as small as we please. The idea that $r/n \to p$ as $n \to \infty$ is thus justified. That $\mathrm{P}' \neq 0$, *i.e.*, it is not certain that we shall ensure $|p - r/n| \leq \varepsilon$ for all n greater than any finite n_0, need not greatly worry us; an infinitesimal uncertainty is always inevitable when we are dealing with chances, however large the number of trials may be (compare, for example, with our belief in the Second Law of Thermodynamics, which is not *certainly* true at any moment).

Cases in practice where n is very large occur in the conceptions in the dynamical theory of gases and statistical mechanics, and here we could almost perhaps identify chance with relative frequency. Generally, however, statements about frequency laws mean rather statements about "laws of chance." The idea of regarding a set of observations as being a random sample from an infinite population is also equivalent to the assumption that each observation is subject to the same system of chances, and obeys a "law of chance." This infinite population is a fiction which it is often, however, convenient to use.

3. *Inverse Probability.*—From certain fundamental axioms it is possible to deduce the standard formulæ and theorems of the calculus of probabilities. These include the theorem of inverse probability, which states that if an event

a can happen through one of n causes $c_1 \ldots c_n$, then the posterior probability of c_s being the true cause after the event a has occurred is

$$\mathrm{P}(c_s \mid a, d) = \frac{\mathrm{P}(a \mid c_s, d) \,.\, \mathrm{P}(c_s \mid d)}{\sum\limits_{r=1}^{n} \mathrm{P}(a \mid c_r, d) \,.\, \mathrm{P}(c_r \mid d)}, \tag{6}$$

where $\mathrm{P}(c_r \mid d)$ is the prior probability of c_r being the true cause on the data d. The terms "posterior" and "prior" need not necessarily have any connection with time, for d may include data obtained at the time of a, or even subsequently. It is, in fact, important to realize that (6) is simply a relation between certain probabilities. If the components of the right-hand side are known, we may deduce $\mathrm{P}(c_s \mid a, d)$; if they are not known, no longer is it obvious that we must evaluate $\mathrm{P}(c_s \mid a, d)$ through (6), or that it will be worth our while to try to do so.

But because it has been sometimes justifiable to regard the $\mathrm{P}(c_r \mid d)$ for various r as equal, and no other method seemed available for inferring which of the causes c_r was most likely, it has been asserted that an explicit use of equation (6) is the only legitimate method for inferring anything about the probable cause of a. Yet Ramsey,* who has stressed that the theorems of probability are simply rules of consistency, has remarked that we may reasonably decide first to assess the posterior probability, based on all our knowledge, and then enquire, if we like, what prior probabilities, based on partial knowledge, this would be consistent with. Something of this sort has indeed unconsciously been done, for taking Fisher's *optimum* value, an estimate of a statistical parameter by a method which will presently be discussed, as identical with the most probable value of that parameter, Haldane† has enquired what assumption about the prior probability of the parameter is sufficient, for consistency with the formula of inverse probability.

Before we proceed further, it will be advisable to write (6) in a form suitable for problems in statistics Let S be a sample of n observations $x_1 \ldots x_n$, and let d denote other relevant knowledge. We assume the observations follow some "law of chance," $f(x, \alpha_i)$, say, where α_i are unknown parameters to be determined—a premiss denoted by f. The possible values of α_i are supposed continuous, so that their true values may be anywhere in certain specified ranges. Then, from (6), the joint probability of particular values of α_i will be given by

$$\mathrm{P}(\alpha_i \mid \mathrm{S}, f, d) = \frac{p(\mathrm{S} \mid f, \alpha_i, d) \,.\, \mathrm{P}(\alpha_i \mid f, d)}{\mathrm{I}\{p(\mathrm{S} \mid f, \alpha_i, d) \,.\, \mathrm{P}(\alpha_i \mid f, d)\}}, \tag{7}$$

---

* "The Foundations of Mathematics, and other logical essays," p. 192 (1931).

† 'Proc. Camb. Phil. Soc.,' vol. 28, p. 55 (1932).

where I in the denominator denotes integration over all possible values of α_i. Since the *chance* of an observation x is independent of what observations have already been obtained (the observations are, of course, assumed uncorrelated), we have further

$$
\begin{aligned}
p(\mathrm{S} \,|\, f, \alpha_i, d) &= p(x_1 \,|\, f, \alpha_i, d) \,.\, p(x_2 \,|\, x_1, f, \alpha_i, d) \\
&= \prod_{r=1}^{n} p(x_r \,|\, f, \alpha_i) \\
&= \prod_{r=1} f(x_r, \alpha_i)\, dx_r
\end{aligned}
\qquad (8)
$$

We may note the peculiar mixture of chances and more subjective probabilities that formula (7) contains. This may explain the confusion that has sometimes arisen as to whether we should be talking about chances or other probabilities. The statistician has been criticized for discussing $p(\mathrm{S} \,|\, f, \alpha_i)$, since it is said he does not know f, α_i, but it is clear from (7) that this chance is essential for any inference about the values of α_i, on " inverse " arguments as well as on direct.

4. Thus Jeffreys' criticism* of Gauss' proof of the normal law would hardly seem to be justified. Gauss showed† that if the arithmetic mean of observations which are assumed to follow a law of error is the most probable value of the true value m, then the error function must be the normal law (this proof is not, of course, a justification of the normal law).

In (7), put $\mathrm{P}(\alpha_i \,|\, \mathrm{S}, f, d) = \mathrm{P}$, $\mathrm{P}(\alpha_i \,|\, f, d) = \mathrm{P}_0$, $\alpha_1 = m$, and take $f(x, \alpha_i)$ to be a function of the error $\varepsilon = x - m$. Then

$$
\mathrm{P} \propto \mathrm{P}_0 \prod_{r=1}^{n} f(x_r - m). \qquad (9)
$$

Take logarithms and differentiate with respect to m. Then P is a maximum for variation in m when

$$
\sum_{r=1}^{n} \frac{\partial}{\partial m} \log f(x_r - m) + \frac{\partial}{\partial m} \log \mathrm{P}_0 = 0, \qquad (10)
$$

and this must be equivalent to

$$
\sum_{r=1}^{n} (x_r - m) = 0 \qquad (11)
$$

---

* *Op cit.*, 70.—*Note.*—I find Dr. Jeffreys is modifying this criticism in a paper not yet out: 'Proc. Camb. Phil. Soc.,' vol. 29, p. 231 (1933). He has, however, kindly let me see his manuscript, and since my discussion includes some points he had missed, it may perhaps be allowed to stand.

† "Theoria motus corporum caelestium," p. 211 (1809).

for all $(x_r - m)$. It follows that P_0 must be independent of m (this condition being necessary because we have demanded to know the *most probable* value of m), and

$$\frac{\partial}{\partial m} \log f(x - m) = \sigma^{-2}(x - m),$$

say, whence

$$f(x - m) \propto e^{-\frac{1}{2}(x-m)^2/\sigma^2}. \tag{12}$$

This proof is quite valid. Jeffreys criticizes the use of $p(S \mid f, \alpha_i)$ and hence equation (9), for he says the probability of an observation x_r, when we do not know the distribution of the probability of error beforehand, must depend on the values already observed, $x_1, \ldots x_{r-1}$; *i.e.*,

$$P(x_r \mid x_1, \ldots x_{r-1}, d) = \phi(x_1, \ldots x_{r-1}), \qquad \text{say.}$$

That is certainly true, but it has no relevance to the proof, for at no stage do the formulæ require this particular probability.

Jeffreys gives an apparent exception in the case where a length is measured by difference. A scale is put arbitrarily against the length, and the ends of the length measured to the nearest unit. The value obtained will be either n or $n + 1$, where the true length is $n + m$, say, where

$$0 \leqslant m < 1.$$

He shows that the chance p_1 of a reading n is $1 - m$, and the chance p_2 of a reading $n + 1$ is m. If we take our origin at n, we have p_1 and p_2 corresponding to readings $x = 0$ and $x = 1$.

The point to notice is that the chances p_1 and p_2 cannot be represented by an error function $f(x - m)$ (unless this has a discontinuous derivative with respect to m). Hence, though Jeffreys shows that the most probable value of m from a number of readings is the mean of the readings, the problem is not strictly relevant to Gauss' proof, since Gauss assumed the unknown function to be a simple function of the error ε. Another case where the unknown m is best determined by the mean, if we do not restrict ourselves to simple error functions, is the Poisson distribution

$$f(x, m) = m^x e^{-m}/x!, \tag{15}$$

where the possible values of x are confined to the positive integers 0, 1, 2, … .

These are all instances of the more general distribution* which can be derived from the condition that the mean shall give the most probable value of the unknown, without the restriction $f(x, m) \equiv f(x - m)$ This is

---

* See Keynes, *op. cit.*, p. 197. It does not seem to have been recognized, however, that $f(x, m)$ is not necessarily a continuous function of x.

$$f(x, m) = \exp\{\phi'(m)(m - x) - \phi(m) + \psi(x)\}. \tag{16}$$

The normal law is the particular case, defined for all values of x,

$$\begin{cases} \phi(m) = -\frac{1}{2}(m/\sigma)^2, \\ \psi(x) = -\frac{1}{2}(x/\sigma)^2 + C \end{cases} \tag{17}$$

The Poisson distribution is the particular case, so defined and non-zero for positive integral values of x only,

$$\begin{cases} \phi(m) = m(1 - \log m), \\ \psi(x) = -\log x\,! \end{cases} \tag{18}$$

The case of measurement by difference, discussed above, since we may write

$$f(x, m) = (1 - m)^{1-x} m^x,$$

where x is restricted to the values 0 and 1, is the particular case, so defined and non-zero for values of x equal to 0 and 1 only,

$$\begin{cases} \phi(m) = (m - 1)\log(1 - m) - m \log m, \\ \psi(x) = 0. \end{cases} \tag{19}$$

5. We saw that the fallacy in the criticism of Gauss' proof seemed to arise from a lack of distinction between the chance of the observations when the distribution of chance is assumed known, and the probability of the observations without this assumption. The failure to distinguish between these two different probabilities would still appear to be causing some confusion.

Thus recently Fisher* has criticized an analysis given by Jeffreys† in connection with prior probabilities, but if we translate into probability notation the relevant steps of Jeffreys' argument, and the parallel argument given by Fisher, we shall see they do not seem to be talking about the same thing.

In (7) we assume f is the normal law, and hence we have only two unknown parameters, $\alpha_1 = m$, the true mean, and $\alpha_2 = 2h^2 = 1/\sigma^2$, where σ is the true standard deviation. We cannot infer m and h from a sample if we attempt to use the inverse probability formula (7), without knowing the prior probabilities of m and h. Jeffreys has tried to show that

$$\mathrm{P}(m, h \mid f, d) \propto dm\, dh/h \propto dm\, d\sigma/\sigma.$$

---

* 'Proc Roy. Soc.,' A, vol. 139, p. 343 (1933).

† 'Proc. Roy. Soc.,' A, vol. 138, p. 49 (1933).

He gives a demonstration of this which rests on the suppositions (i) that the probability that the third of three observations shall lie between the other two, when m and h are unknown, is $\frac{1}{3}$, *i.e.*,

$$P(\{x_3\} \mid x_1, x_2, f, d) = \tfrac{1}{3} P([x_3] \mid x_1, x_2, f, d), \tag{20}$$

where $\{x_3\}$ denotes the fact that x_3 lies in the interval (x_1, x_2), $[x_3]$ that it lies in the entire range $(-\infty, \infty)$; and (ii) that

$$P(m, h \mid f, d) \propto F(h)\, dm\, dh, \text{ say.}$$

He writes further $x_1, x_2 = \pm a$, and proceeds to evaluate each side of (20) by means of inverse probability, considering first $P(x_3 \mid x_1, x_2, f, d)$.

We have

$$p(x \mid f, m, h) \propto h e^{-h^2 (x-m)^2}\, dx;$$

$$P(x_3, m, h \mid x_1, x_2, f, d) = P(m, h \mid x_1, x_2, f, d)\, p(x_3 \mid f, m, h)$$

Further, by the formula of inverse probability (7),

$$P(m, h \mid x_1, x_2, f, d) \propto p(x_1, x_2 \mid f, m, h)\, P(m, h \mid f, d),$$

since the denominator of the formula will be a function of x_1 and x_2 only, that is a function of a only, and can be ignored. Hence

$$P(x_3, m, h \mid x_1, x_2, f, d) \propto h^3 F(h)\, e^{-2h^2(m^2+a^2) - h^2(x_3 - m)^2}\, dm\, dh\, dx_3$$

To find $P(x_3 \mid x_1, x_2, f, d)$, we have to integrate with respect to m and h, and to evaluate each side of (20), we have to integrate further with respect to x_3 from $-a$ to a, or from $-\infty$ to ∞. Integrating with respect to m and x_3, Jeffreys puts (20) in the form

$$\int_0^\infty h\, F(h)\, e^{-2h^2a^2} \operatorname{erf}\{\sqrt{(\tfrac{2}{3})}\, ha\}\, dh = \tfrac{1}{3} \int_0^\infty h\, F(h)\, e^{-2h^2a^2}\, dh,$$

and if this is to hold for all a, he shows the only possible form for $F(h)$ is

$$F(h)\, dh \propto dh/h \propto d\sigma/\sigma \tag{21}$$

Fisher criticizes this result, but it is possible that he has misinterpreted Jeffreys' work, for he uses direct arguments, and considers no probabilities other than chances. Thus, when commenting on Jeffreys' supposition (i), he says "*for any particular population*, the probability will generally be larger when the first two observations are far apart than when they are near together." I have italicized the first words to show that he is not considering the probabilities in (20) on the same data as Jeffreys, but on the data f, m, h. He thus considers $p(\{x_3\} \mid f, m, h)$. Since, given m and h, the chance that x_3 should fall within the interval (x_1, x_2) clearly depends on the values of x_1 and x_2, we may write

$$p(\{x_3\} | f, m, h) = \psi(x_1, x_2), \text{ say.} \tag{22}$$

Fisher's work essentially consists of finding the value of $\psi(x_1, x_2)$ when averaged for all possible values of x_1 and x_2. He writes $x_1, x_2 = u \pm v$, and $x_3 = u + c$, and thus

$$p(x_1, x_2, x_3 | f, m, h) = (h/\sqrt{\pi})^3 e^{-h^2 \Sigma (x_r - m)^2} dx_1 \, dx_2 \, dx_3$$

is equivalent to

$$p(u, v, c | f, m, h) = 2(h/\sqrt{\pi})^3 e^{-2h^2[(u-m)^2 + v^2] - h^2(u-m+c)^2} du \, dv \, dc.$$

Integrating for u from $-\infty$ to ∞, then for c from $-v$ to v or from $-\infty$ to ∞, then for v from $-\infty$ to ∞, Fisher finds the average value of $p(\{x_3\} | f, m, h)$ is $\frac{1}{3}$, and the average value of $p([x_3] | f, m, h)$ is 1 (the latter result being obvious).

Fisher inserted the factor $\mathrm{F}(h) \, dh$, representing the prior probability of h, in his work, but clearly this is irrelevant to his argument, just as he found the prior probability of m also irrelevant; for all the time he is using direct arguments His result would therefore seem to have little connection with Jeffreys', and equations (20) and (22) are two entirely different propositions.

Jeffreys' result depends on the validity of equation (20), and the particular criticism that this invites is that it seems to have been justified on a peculiar inversion of statement. For he says " the law says nothing about the order of occurrence of errors of different amounts, and therefore the middle one in magnitude is equally likely to be the first, second, or third made (provided, of course, that we know nothing about the probable range of error already)." But even if we know all about the distribution of error beforehand, the three observations will be different in value, and it is still equally likely that the middle observation should happen to be the first, second, or third made. Thus, whether we know the range of error or not, the probability that the *middle* one in magnitude of three observations was the *third* one made is $\frac{1}{3}$. But this tells us nothing about the probability that, given two observations, a *third* observation will be the *middle* one; as we have seen, if we know the range of error, this probability will be a function of the values of the first two observations, if we do not know the range of error, its value is still undetermined.*

---

* Dr. Jeffreys agrees with me that his argument, as he has originally given it, is perhaps hardly sufficient. If I interpret him correctly (using my notation), he suggests rather the following:

The principle is that if the range of error is completely unknown, $\mathrm{P}(\{x_3\} | x_1, x_2, f, d)$ is independent of x_1 and x_2 (and in particular of $x_1 - x_2$). Then the prior probabilities of

6. *Criticism of the Explicit Use of the Theorem of Inverse Probability.*—The particular criticism given above, is, however, of less importance than criticism on more general lines, especially as Jeffreys has given a former argument for the validity of (21)—that it makes the probability of a value of σ or h in any range independent of the scale of measurement.

The attempt to use the formula of inverse probability for the purpose of theory inevitably means that we cannot be concerned with finding the real prior probabilities in relation to particular problems, even if that were possible, but merely with finding some simple mathematical function which when put in place of these unknown probabilities will make our work appear reasonably consistent.

This substitution of a simple function for a prior probability, which if it could be evaluated at all would certainly needs all the data d to be enunciated, gives the posterior probability in an exact form which is highly misleading. Moreover, the formula in trying to make our final inference about a parameter for us and to give us the exact probability of every possible value, is compelled to mix the information which can be got out of a sample with what other knowledge or notions we may have. This indicates that we are asking too final a question ; and in practice it would mean that only when we could ignore our previous knowledge as negligible or irrelevant could we hope to make any headway at all.

We should remember there are other points still to be considered : thus the premiss f of, say, the normal law is still left in the data, and if we really want the probability of possible values of such a parameter as the true mean, we have to consider all possible " laws of chance," of which the normal law—though having, it is true, a reasonable prior probability—is only one. Anyone formally using the inverse probability formula has still to consider the specific practical problem, whether the sample is representative, whether conditions

---

the orders (with respect to magnitude) $(x_1x_2x_3)$, $(x_2x_1x_3)$, $(x_1x_3x_2)$, $(x_3x_1x_2)$, $(x_2x_3x_1)$, $(x_3x_2x_1)$, are all equal. And if $x_1 < x_2$, say (already known by observation),

$$
\begin{array}{ll}
P(x_3 < x_1 | x_1, x_2, f, d) \text{ is the case } & (x_3x_1x_2), \\
P(\{x_3\} | x_1, x_2, f, d) & (x_1x_3x_2), \\
P(x_2 < x_3 | x_1, x_2, f, d) & (x_1x_2x_1).
\end{array}
$$

As these alternatives are exhaustive and exclusive, the probabilities are each $\frac{1}{3}$.

This means, I think, that if we were completely ignorant of the range of error beforehand, the apparent magnitude of $x_1 - x_2$, since it depends on what scale we have adopted, cannot give us any information whether x_3 is likely to fall within the interval (x_1, x_2) or not. It is thus equivalent to saying that the probability of a value of h in any range should be independent of the scale of measurement.

have altered, whether the normal law, say, holds, and any advantage the explicit use of the formula appeared to give him, in obtaining an exact probability for each possible value of an unknown, disappears.

All this, however, is of minor importance compared with the fundamental objection to a method which requires the evaluation of prior probabilities However necessary it may be to bear in mind prior probabilities when we are making any inductive inference from a sample, no rules can ever be given for their evaluation. For if they could, we should always be able to *deduce* our induction as an inference on certain data with a unique probability: this identification in terms of probability of the logically distinct concepts of deduction and induction is one that very few, at any rate, are prepared to recognize

It would, therefore, hardly seem desirable to persevere in the use of a theory which incorporates such indefinite quantities if we can devise any other methods which will help us to make inferences.

7 *Arguments about Chances.*—If we decide to obtain first what information we can simply from the sample, we naturally consider

$$p\,(\mathrm{S}\,|\,f,\,\alpha_i) = \prod_{r=1}^{n} f\,(x_r,\,\alpha_i)\,dx_r \tag{23}$$

This, as we have seen, is not all that we require in the inverse probability formula, and we no longer try to develop a formal theory to find the exact posterior probability of a parameter, instead, we may start from (23), arguing legitimately about chances, and hence obtain other methods of making inferences about the parameters.

These arguments about chances have been largely developed by Fisher. The method of maximizing the chance in (23) is known as the "method of maximum likelihood," and the value of a parameter α_i so obtained Fisher calls the *optimum* value.* This value thus gives the highest chance to the observations we actually have. For example, for the normal law,

$$p\,(\mathrm{S}\,|\,f,\,m,\,\sigma) \propto \sigma^{-n}\,e^{-\Sigma(x_r-m)^2/2\sigma^2}. \tag{24}$$

Taking logarithms, and differentiating with respect to m and σ we find the *optimum* values of m and σ are given by

$$\begin{cases}\Sigma\,(x_r - m)/\sigma = 0,\\ \Sigma\,(x_r - m)^2/\sigma^3 - n/\sigma = 0,\end{cases}$$

---

* 'Phil. Trans.,' (*loc. cit.*), p. 327.

whence

$$\begin{cases} m = \Sigma x_r/n = \bar{x}, & (25) \\ \sigma^2 = \Sigma\,(x_r - m)^2/n. & (26) \end{cases}$$

If we had assumed, with Jeffreys, that the prior probability of σ varies as $1/\sigma$, we should have found, using the formula of inverse probability, that the most probable value of σ was apparently given by

$$\Sigma\,(x_r - m)^2/\sigma^3 - n/\sigma - 1/\sigma = 0,$$

whence

$$\sigma^2 = \Sigma\,(x_r - m)^2/(n + 1) \tag{27}$$

The result (26) suggests a reasonable value to take for σ, and no one appears to prefer the result in (27)

We should notice that (26) requires a knowledge of m, and if m is unknown, it is for most purposes more satisfactory to obtain an estimate of σ^2 which is uncorrelated with m; otherwise, a bias is introduced when we substitute for m its estimate from the sample, $\bar{x}$ Starting from (23), we can investigate the joint chance* of $\bar{x} - m$, and $\Sigma\,(x_r - m)^2$, but we find we are obliged to consider rather the chance of a value $\Sigma\,(x_r - \bar{x})^2 = \Sigma$, say, this being independent of the chance of a value $\bar{x} - m$, and given by

$$p\,(\Sigma \mid f, m, \sigma)\,\alpha\,\Sigma^{\frac{1}{2}(n-3)}\,e^{-\frac{1}{2}\Sigma/\sigma^2}\,d\,\Sigma/\sigma^{n-1}. \tag{28}$$

Since this chance is independent of m, we can find an *optimum* value of σ which is independent of m by finding what value of σ gives the highest chance to Σ. Taking logarithms and differentiating with respect to σ, we obtain

$$\Sigma/\sigma^3 - (n - 1)/\sigma = 0,$$

whence

$$\sigma^2 = \Sigma/(n - 1). \tag{29}$$

This estimate of σ^2 is denoted by s^2. Since the chance distribution of $\Sigma\,(x_r - m)^2$ alone is that given in (28), but with $n - 1$ replaced by n, we see that the substitution of $\bar{x}$ for m has the effect of diminishing the number of degrees of freedom of Σ from n to $n - 1$, and thus where our original estimate of σ^2 is that in (26), the substitution of $\bar{x}$ for m makes the corresponding estimate in (29) more satisfactory than the estimate in (26) with $\bar{x}$ written for m. This is especially true when n is small, or when variances of several samples are considered, since the mean value of s^2 is σ^2.

---

* Wishart and Bartlett, 'Proc. Camb. Phil. Soc.,' vol. 29, p. 260 (1933).

There is a modification of a similar kind for two variables x and y; though the *optimum* value of ρ, the correlation coefficient, when estimated simultaneously with m_x, m_y, σ_x, σ_y, is the usual formula,

$$r = \Sigma (x_r - \bar{x})(y_r - \bar{y})/\{\Sigma (x_r - \bar{x})^2 \Sigma (y_r - \bar{y})^2\}^{\frac{1}{2}}, \tag{30}$$

the *optimum* value of ρ when estimated singly from the chance distribution[*] of r is given by $r - r(1 - r^2)/2(n-1) + 0(n^{-2})$ Fisher has suggested the change of variable,

$$\begin{cases} \rho = \tanh \zeta, \\ r = \tanh z, \end{cases}$$

since the distribution of z, unlike that of r, is very nearly normal. Whereas the distribution of z has, however, a slight bias, its mean value being

$$\zeta + \rho/2(n-1) + 0(n^{-2}),$$

the *optimum* value of ζ which corresponds to the modified estimate of ρ is

$$z - r/2(n-1) + 0(n^{-2}), \tag{31}$$

and it is easy to see that the bias has been corrected to the first order. Equation (30) gives the simpler estimate to use, the bias in the mean value of z that it introduces being unimportant for reasonably sized n unless a pooled estimate from several samples is to be made, when the estimate of ζ corresponding to the final estimate of ρ should be obtained not from the various values of z, but from the corrected values given in (31).

It is not asserted that the *optimum* value of a parameter is necessarily the most probable value. If we already knew or suspected the true value α of a parameter before the sample was taken, we should be interested then not in the *optimum* value a as the most probable value, but in the chance that a discrepancy at least as great as that between the true value and the *optimum* value should have arisen. This suggests further arguments about chance which may be regarded as complementary to the "method of maximum likelihood," in the problems of testing significance, and assessing limits within which the true value of a parameter is likely to lie. Thus we may test the hypothesis that the parameter has a specified value α_0 by finding the chance of a discrepancy greater or equal to $|a - \alpha_0|$. Similarly, by finding values α_1 and α_2 ($\alpha_1 < a < \alpha_2$, where α_1, α_2, and a are here assumed all positive), such that the chance of a discrepancy $a - \alpha_1$ or greater, or $\alpha_2 - a$ or greater, is some chosen small number, *e g.*, $p = \cdot 05$, we can infer that it is likely that

---

[*] See Fisher, 'Biometrika,' vol. 10, p. 507 (1915); 'Metron,' vol. 1, p. 1 (1921).

$$\alpha_1 \leqslant \alpha \leqslant \alpha_2,$$

and hence assign probable limits to the value α. This method Fisher has called the method of fiducial probability.*

We see that these methods involve a knowledge of the "laws of chance," or chance distributions, of particular functions of the observations which are our estimates of the unknown parameters, *e.g* , the estimate of variance and the estimate of the correlation coefficient. The investigation of these chance distributions has therefore been one of the most important parts of theoretical research.

8 *The Relations of the "Optimum," and the "Most Probable Value," to the Sample.*—In the investigation of Gauss' proof of the normal law, we saw that in order to examine the postulate that the arithmetic mean was the *most probable* value, we had to make an assumption about the prior probability of the true value m. It is rather absurd that in order to find the relation between the arithmetic mean and the properties of the normal law of error, such an assumption should be necessary, and if we had asked the more reasonable question—what law of error makes the mean most representative of a sample of observations ?—the question of prior probability would not have arisen.

This fundamental difference between the most probable value, and the *optimum* value, which represents the information in the sample, is illustrated by cases where the true value is known. If, for example, a student of physics made a number of determinations of J, the mechanical equivalent of heat, he would expect, although he knows the most probable value of J is $4 \cdot 18 \times 10^7$ ergs, to be able to give from his experiments *his* estimate, what we may call his *optimum* value.

Again, if we consider the problem of observations made over a number of years in two variates which are correlated, we may want to know whether conditions have changed, whether, if we divide the series of observations at its mid-point, the correlations for the two groups are the same. We may investigate this question by taking the *optimum* values for the correlations of the two groups, and enquiring whether they are significantly different. On the other hand, if we tried to find the most probable values of the two correlations, naturally the most probable value of the correlation of the second group depends partly on the information in the second group and partly on the probability that the correlation is the same as that of the first group, and that again depends partly on the answer to the question we are asking.

---

* 'Proc. Camb. Phil. Soc ,' vol. 26, p. 528 (1930)

The method of maximum likelihood consists in finding the value of a parameter which gives the highest chance to the observations we have, or to the function of the observations which is our estimate of the parameter It therefore considers simply the sample we have, and obtains an estimate which best represents that sample. Thus Fisher has shown that this estimate is "efficient," that is, the variance of its chance distribution is as small as possible.*

In the particular but common case where the sample is the only relevant information we have, it is natural to take the *optimum* value as our best estimate of a parameter This has been confused with a use of the inverse probability formula, although the *optimum* value a when estimated from the chance distribution

$$p = g(a, \alpha)\, da,$$

say, is invariant for change of variable, whereas not only the mean and the mode† of this distribution, but also the most probable value of α—on the assumption that nothing being known about α, any value has an equal prior probability—all vary.

Jeffreys,‡ in criticizing these methods, says "the whole reason for attaching any importance to Fisher's 'likelihood' is that it is proportional to the posterior probability given by Laplace's theory, and it has no meaning outside the original sample except in terms of this theory."

But nothing obtained from the sample has any meaning outside the sample unless we give it one, and what meaning we give it must rest with our own judgment If we decide to start simply with

$$p(S \mid f, \alpha_i),$$

we may, as Ramsey has suggested,§ define the process of maximizing this expression in order to find estimates of parameters as a process in pure mathematics, and consequently entirely distinct from a use of the formula of inverse probability.

---

* 'Proc. Camb. Phil. Soc.,' vol. 22, p. 600 (1925).

† *I.e.*, the most probable value of a for given α, not to be confused with the most probable value for α for given a.

‡ 'Proc. Camb. Phil. Soc.,' vol. 29, p. 83 (1933).

§ *Op. cit.*, p. 204.

It is obviously true that inverse probability (if we choose to associate this term with the whole idea of making inductions and inferring causes from events) must have some place in our inferences. We may observe, for example, that the posterior probability of a correlation being significant should be regarded as zero, if the prior probability was believed to be zero; but presumably no one would ever calculate a correlation coefficient if he did not believe it had some significance. Inverse probability, in its widest sense, is implicitly contained in an inductive inference in any science, but that is no reason why, because the particular probabilities known as chances form a large part of the subject-matter of statistics, it should be used explicitly, and entangle these chances with prior probabilities and previous or *a priori* knowledge.

I am indebted to Dr. J. Wishart for advice and criticism during the preparation of this paper

Summary.

Frequency laws in statistics mean "laws of chance." The distinction between these chances and probabilities on other data is emphasized; especially as it has not invariably been perceived.

An attempt is made to show why exact arguments about chance are more fundamental as a mathematical basis for statistical theory and inference than the formula of inverse probability, which, in statistics, is a hybrid of exact chance, and unknown prior probabilities. Any attempt to evaluate the latter for the purpose of formal theory can lead only to misleading assumptions.

Examples of arguments about chances are Fisher's methods of maximum likelihood and fiducial probability.

Such arguments do not pretend to abolish the judgment and common sense necessary when we use general theory to help us make, from a particular sample, an inference about the "population"

Experiments on the Raman Effect at very Low Temperatures.

By G. B. B. M. SUTHERLAND, Trinity College, Cambridge

(Communicated by T. M. Lowry, F.R.S.—Received March 16, 1933.)

Introduction

Hitherto, the experiments which have been carried out on the Raman effect at low temperatures have been performed on substances of which the Raman spectra were unobtainable, or only obtainable with difficulty, at ordinary temperatures. Thus Daure* has obtained the Raman spectra of liquid methane, ethylene, ethane, propane, and ammonia, while McLennan and his co-workers† have studied those of liquid hydrogen, oxygen, nitrogen, methane, helium, nitrous oxide and solid carbon dioxide. The primary object in these researches was to obtain the spectrum of the substance, and the influence of temperature on the spectrum has received very little attention. While the present work began as an attempt to obtain the Raman spectrum of nitrogen tetroxide, it has become increasingly evident in the course of it that the study of Raman spectra at very low temperatures may well prove to be a very fruitful field of research.

It is clear that there will be at least two distinct possible effects of a change in temperature of the scattering substance on its Raman spectrum. If the change in temperature causes a change in the molecular structure of the substance (such as association of simple molecules into more complex aggregates, or a change in the crystalline form of the substance), then one may expect the appearance of new lines in the spectrum. Secondly, a change in temperature will result in a new distribution of the molecules in the various rotational and vibrational energy levels; the effect of this will be to alter the character of the individual lines and bands in the spectrum without, however, giving rise to any new lines. The following paper gives a description of a new and very simple apparatus for the observation of Raman spectra at low temperatures, together with the results and discussion of the preliminary experiments on a few simple substances.

---

* 'Ann. Physique,' vol. 12, p. 375 (1929).

† McLennan and McLeod, 'Trans. Roy. Soc. Canada,' vol. 22, p. 413 (1928); McLennan, Smith and Wilhelm, *ibid.*, vol. 23, p. 19 (1929), *ibid.*, vol. 23, p. 247 (1929), *ibid.*, vol. 24, p. 197 (1930); McLennan and Smith, 'Canadian J. Res.,' vol. 7, p. 551 (1932).

Experimental.

The Dewar flask A was specially constructed from two Pyrex glass tubes which were sealed together at B and closed at the lower ends DD by flat plates of Pyrex glass. The outer tube was drawn down as indicated, and by successively breaking it off and rejoining it to a narrower tube a few projecting rings CC could be made which served as stops in partially collimating the scattered light. The whole flask was silvered inside excepting the plane windows DD, and a strip EE, 2 cm. wide, running vertically up one side. The stops CC were blackened with a suitable non-reflecting ink. The small aluminium stand F fits snugly into the flask, and the top of it is on a level with the end of the clear strip EE. The observation tube G is supported on F; its upper end may be held lightly in a clamp to keep it vertical, but if F is well designed this is not essential.

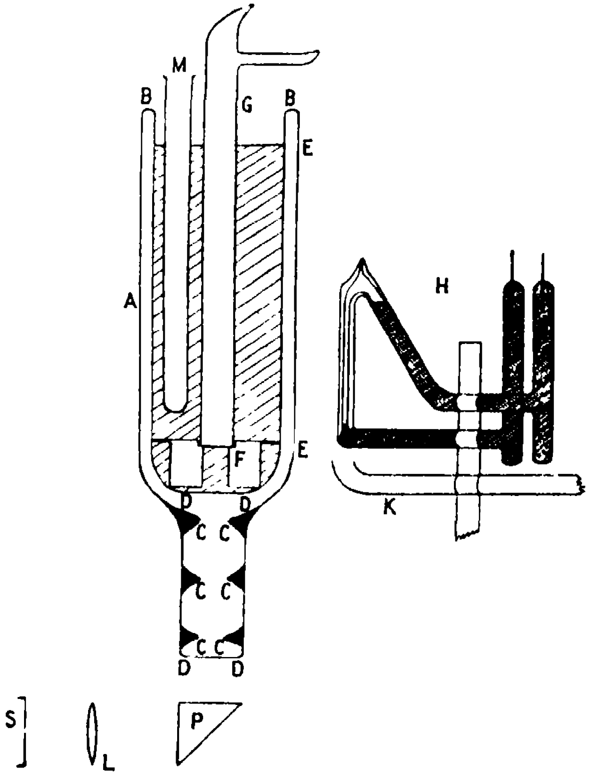


Fig. 1.

The quartz mercury arc H is placed parallel, and opposite to the clear strip

EE, and 4 to 5 cm. away from the flask. This type of arc runs very satisfactorily in the vertical position. It was made from 3 mm. quartz capillary tubing bent as shown, and was found to be a very powerful source of the radiation at 4047 A. and 4358 A. when operating with a potential drop of 120 volts across the electrodes and carrying a current of approximately 1 ampere The blast of air from the brass tube K served to keep both the lamp and the side of the Dewar flask from becoming too hot The scattered light was reflected into the slit of the spectrograph S by the right angle prism P, the sphero-cylindrical lens L being used to focus and concentrate it The Dewar flask was supported in a wooden stand, and once the alignment of the parts had been secured so that only light which came vertically down the centre of the flask was focussed into S, then it never required doing again, for the observation tube could be removed and always put back in the same position

Although this arrangement was arrived at in order to study conveniently the Raman spectra of many substances at liquid air temperatures, a fairly continuous range of temperatures down to —100° C may be obtained if, instead of filling the flask with liquid air, one fills it with ethyl alcohol. The alcohol bath is then cooled to the desired temperature by pouring sufficient liquid air into the pyrex tube M which dips into it

The arrangement described possesses distinct advantages over the earlier ones. The use of an almost completely silvered Dewar flask instead of the unsilvered ones used previously provides better heat insulation, and so requires less liquid air for its operation This can be quite an important factor in making long exposures when the flask may have to be left unattended overnight. The reflections from the silvering on the inner walls also help to intensify the exciting light, giving the " light furnace " effect so often used in experiments on Raman spectra. The shielding of the spectrograph from any direct light is more complete than in most of the earlier arrangements. The advantage of taking the light which is scattered vertically downwards instead of that which is scattered upwards is that it occasionally happens that a substance may scatter a certain frequency quite strongly when in the solid form at very low temperatures, but that its uncondensed vapour absorbs the same frequency. When this happens the slightest trace of uncondensed vapour in the upper part of the observation tube may cause complete absorption of the scattered light

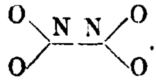
In most of the work to be described, liquid air was used as the cooling agent, but a few photographs were taken at temperatures ranging from —70° C to

20° C. The substances examined were nitrogen tetroxide, ozone, ice, methane, ammonia, carbon tetrachloride, and liquid air. For each substance the exciting lines from the mercury arc were those at 4047 A. and 4358 A., although lines were occasionally observed scattered by the strong group near 3650 A. An attempt to use a similar apparatus in quartz and employ the strong mercury line at 2537 A has not yet been successful, as small traces of oil in the liquid air seem to absorb this line so much that the resulting scattering of it is too weak to give observable Raman lines The wave-lengths of the lines in the Raman spectrum were obtained from a Hartmann interpolation formula using the mercury lines as standards. The spectrograph used was a two-prism glass instrument with an aperture of f/16 and having a mean dispersion of 18A per mm. near 4400A The plates used were " Gevaert Super Chromosa "; these were found to be most satisfactory from the standpoints of speed, fineness of grain, and freedom from halation.

Nitrogen Tetroxide.

This substance was the first to be examined, as a knowledge of its complete Raman spectrum formed an essential complement to work which was being done by the author on its infra-red absorption spectrum * At ordinary temperatures nitrogen tetroxide is a gas of which about 20% is dissociated into the dioxide On cooling it the proportion of dissociated molecules steadily decreases, it liquefies as a clear reddish-brown liquid, the colour being due to the dissociated molecules. As the liquid is further cooled, the colour gradually disappears until on solidification at —10° C. it forms either a transparent or a perfectly white crystalline mass, which is composed entirely of undissociated molecules.† The transparent form is rather unstable, and in all the work described here it was the latter form which was employed ‡

The Raman scattering of nitrogen tetroxide has been previously studied by Menzies and Pringle§ at a temperature of —80° C. They reported one line only at 275 cm. $^{-1}$, but it was felt that many more must exist Table I, which summarizes the results obtained, shows that it has a very rich spectrum.

The line at 283 cm.$^{-1}$, which obviously corresponds to the one observed by Menzies and Pringle, is by far the strongest line in the spectrum. Although exposures of several hours were necessary to bring out the other lines, this one could be obtained in a few minutes. This line was later examined under higher dispersion and found to be a doublet of which the components seemed to be of equal intensity and had a separation of 3 cm^{-1} The interpretation of this line and of the lines at 500 cm.$^{-1}$, 813 cm $^{-1}$, 1337 cm.$^{-1}$, 1382 cm. $^{-1}$, and 1724 cm.$^{-1}$ has been given elsewhere.* It is sufficient to remark here that these have all been satisfactorily explained in terms of the fundamental frequencies of a molecule of the form . Vegard † has proposed from his results on X-ray powder photographs that the crystal lattice of solid nitrogen peroxide is built up from NO_2 molecules each of which is linear and symmetrical. These proposals have been criticized by Hendricks,‡ who claims that an unambiguous determination of the structure is not possible from Vegard's photographs. From the above results on the Raman spectrum we may at once dismiss Vegard's theory as quite untenable. If the crystal were composed of linear and symmetrical NO_2 molecules it would exhibit only one strong Raman frequency which would lie in the neighbourhood of 1300 cm.$^{-1}$. The greatest number of Raman frequencies it could possibly have would be two, *i.e.*, a splitting of this single frequency into two by accidental resonance degeneracy such as occurs in carbon dioxide.§ The weak third and fourth lines which appear in carbon dioxide would not be observable as the number of molecules in the first vibrational level of such a linear molecule at the temperature of liquid air would be infinitesimally small.

---

* Sutherland, p. 342.

† Mellor, " Treatise on Inorganic Chemistry," vol. 8.

‡ The sample of nitrogen peroxide used was provided by the Research Department of Nobel's Explosive Company, to which the author wishes to express his indebtedness.

§ 'Nature,' vol. 127, p. 707 (1931).

Table I —Raman Spectrum of Solid Nitrogen Tetroxide at —190° C.

Exciting frequency	Raman frequencies in cm.$^{-1}$.								
22938	28†	57	79	283	500	812	1336	1382	1724
24705	—	56	80	284	*	814	1338	1382	*
Intensity	medium		strong	very strong	med.	strong	medium		

* Obscured by Mercury line
† Observed only as an anti-Stokes line

---

* Sutherland, *loc. cit.*
† 'Z Physik,' vol. 68, p. 184 (1931); vol. 71, p. 299 (1931).
‡ 'Z. Physik,' vol. 70, p. 699 (1931).
§ Fermi, 'Z. Physik,' vol. 71, p. 250 (1931).

The lines at 28 cm.$^{-1}$, 57 cm.$^{-1}$ and 79 cm.$^{-1}$ are more difficult to interpret. The possibility that these are fundamental vibration frequencies may be immediately ruled out on account of their magnitude, and in view of the fact that an application of Placzek's rules* shows that not more than six of the twelve fundamental frequencies of this molecule may appear in the Raman effect. The separations between them are obviously much too great for them to represent pure rotation frequencies, supposing that rotation were possible in this crystal. It seems not unreasonable, however, that these may represent the "oscillational frequencies" proposed by Pauling† in which the molecule oscillates about a mean position in the crystal without ever having enough energy to make a complete rotation. We note that these frequencies form a linear progression, and recall that Pauling has shown that the energy levels corresponding to such motions will be given to a first approximation by the familiar formula

$$E = (n + 1)\, h\nu_0 \qquad n = 0, 1, 2, 3, \ldots$$

In this connection it is important to add that several exposures were made of nitrogen tetroxide at various temperatures between −20° C. and −70° C., when these lines were again observed, and also faint lines at approximately 105 cm.$^{-1}$, 125 cm.$^{-1}$ and 142 cm.$^{-1}$ from the mercury line‡ at 4358 A. These are probably the next three members of the series, but further experiments under higher dispersion and with a more suitable exciting line will be required to determine their positions accurately.

If this interpretation of these lines is correct, it is of interest to estimate whether rotation might be possible in this crystal. Pauling's criterion is that the transition from oscillation to rotation will take place for molecules in the nth vibrational state, where n is given by the equation

$$(n + 1) = \frac{\beta\nu_0}{4\Theta}.$$

Here

$$\beta = \frac{h}{K} \text{ and } \Theta = \frac{h^2}{8\pi^2 IK}.$$

Now the moment of inertia I is unknown, but we may safely estimate that a mean value for it will be not less than 2×10^{-38} gm. cm.^{2}. This leads to a value of 49·6 for n in the above equation and corresponds to an energy level so high that the fraction of molecules possessing this energy even at the melting point (—10° C.) is quite negligible.

---

* 'Leipziger Vorträge,' p. 71 (1931).

† 'Phys. Rev.,' vol. 36, p. 430 (1930).

‡ The line at 142 cm.$^{-1}$ was observed only as an anti-Stokes line, since the Stokes line coincides with a mercury line.

Ozone.

In this part of the work the author had the assistance of Dr. S L. Gerhard of the University of Michigan, who was engaged at that time on an investigation of the infra-red absorption spectrum of ozone. A preliminary account of the results has already appeared,* but a more complete discussion of them can now be given. For details of the preparation of the ozone, reference should be made to Gerhard's paper.† Pure liquid ozone is a very dark purple liquid, which may explode violently when under low pressures It was considered safer to work with a 30 % solution of ozone in liquid oxygen ‡. The top of the observation tube was always left open to the atmosphere and no explosions occurred in the course of the experiments. All the work was done at liquid air temperatures.

This solution of ozone was found to scatter the mercury line at 4047 A. fairly strongly, but seemed to absorb the line at 4358 A and all mercury lines of longer wave-length to a considerable extent. Exposures were made lasting as long as 25, 40, 60, and 80 hours, but no definite Raman lines were observed, although the Rayleigh scattering of the line at 4047 A. seemed to be as great as it had been with nitrogen tetroxide, for comparable times of exposure On two plates there were indications of what looked like a doublet corresponding to a mean frequency shift of 1280 cm $^{-1}$, but these could not be definitely confirmed as there are some weak mercury lines in this neighbourhood. In view of Gerhard's analysis of the infra-red spectrum and his assignment of the fundamental frequencies at 528 cm.$^{-1}$, 1033 cm.$^{-1}$, and 1355 cm.$^{-1}$ it seems probable that the observed doublet is spurious

The failure to observe any strong Raman spectrum from ozone may be due to two causes. The conditions for observation may have been unfavourable, or the Raman scattering may at best be extremely weak. From the fact that the Rayleigh scattering of the line at 4047 A. was so strong the conditions would seem to have been favourable, yet it must be remembered that the ozone solution did seem to absorb very strongly all light of wave-length longer than 4358 A , which means all possible Raman lines of frequency greater than 1750 cm $^{-1}$, excited by the line at 4047 A. The absorption of ozone may continue for some way towards shorter wave-lengths than 4358 A., so that there is the possibility that potential Raman frequencies lower than 1750 cm.$^{-1}$ were reabsorbed by the liquid. It is more probable that the weakness of the Raman spectrum is due to more fundamental causes It would indicate that the ozone molecule does not possess a highly symmetrical form such as might be represented by an equilateral triangle or a stright line model. A molecule having either of these forms would possess one symmetrical mode of vibration* which involves large changes in the polarizability and would therefore appear with great intensity in the Raman effect.† The conclusion that the form of the ozone molecule is triangular but not equilateral was also that reached by Gerhard from his investigations of the infra-red spectrum.

---

* Sutherland and Gerhard, 'Nature,' vol. 130, p. 241 (1932).

† 'Phys. Rev.,' vol. 42, p. 662 (1932).

‡ A few exposures were also made using pure ozone

Ice.

The Raman spectrum of ice at 0° C. has been already obtained by Rao‡ and by Ganesan and Venkateswaran,§ who observed three very diffuse lines at 3193 cm.$^{-1}$, 3391 cm.$^{-1}$, and 3549 cm $^{-1}$. These are to be compared with the maxima in the broad diffuse water band at 3216 cm.$^{-1}$, 3435 cm.$^{-1}$, and 3582 cm.$^{-1}$, obtained by several observers,‖ and with the line reported for water vapour at 3654 cm.$^{-1}$ by Johnston and Walker.¶ It has been suggested by Rao** that the appearance of three bands in water and ice is due to there being three different molecules viz, H_2O, $(H_2O)_2$, and $(H_2O)_3$. As the band at 3193 cm $^{-1}$ becomes extremely weak at higher temperatures (*circa* 80° C.) it is associated with the $(H_2O)_3$ molecules, while the band at 3549 cm.$^{-1}$, which increases slightly in intensity with increase in temperature, is associated with the H_2O molecule. If this explanation is correct, then one would expect little change in the Raman spectrum of ice if photographed at the temperature of liquid air instead of at 0° C., since the association of simple molecules into larger aggregates with decrease in temperature, while quite probable in the liquid, is rather improbable in the solid

---

* For the linear model it would be the mode of vibration in which the two outer atoms move symmetrically towards and away from the centre one; for the equilateral one it would be that in which the atoms move along the medians of the triangle.

† Placzek,' 'Z. Physik,' vol. 70, p. 84 (1931).

‡ 'Ind J. Phys.,' vol. 3, p. 105 (1928).

§ *Ibid.*, vol. 4, p. 196 (1929).

‖ Kohlrausch, "Smekal-Raman Effekt" (1931).

¶ 'Phys. Rev.,' vol. 39, p. 535 (1932).

** 'Proc. Roy. Soc.,' A, vol. 130, p. 409 (1930).

Accordingly, a few exposures were made of ice at the temperature of liquid air. The spectrum was found to consist of one quite intense and fairly sharp line at 3090 cm.$^{-1}$ with a faint companion at 3135 cm.$^{-1}$. Such a great change, both in the frequency and in the character of the Raman radiation is not easily explained on Rao's theory It might be supposed that a gradual polymerization into $(H_2O)_3$ molecules, or even larger aggregates, takes place in the crystal as the temperature decreases, but this is in direct contradiction to the X-ray observations of Barnes,* who found no change in the crystal structure of ice between 0° C. and $-183°$ C. Moreover, the work of Dennison† has pointed to the conclusion that the molecules in ice are very probably of the form $(H_2O)_2$

The most striking fact is the great increase in the sharpness of the ice lines as the temperature is lowered. This was also noticed (although to a much smaller extent) in the case of nitrogen tetroxide, for which the lines obtained at $-190°$ C were appreciably sharper than those taken at temperatures of $-30°$ C. to $-50°$ C. This would seem to indicate that the diffuseness is due largely to strong, variable, intermolecular fields present in water and in the crystals. As the temperature is lowered the variations in these internal fields due to the random heat motions of the molecules become less and less, so the possible changes produced in the vibration frequency become less and less and the line becomes sharp and well defined instead of diffuse and broad The presence of such intermolecular fields would also explain the large changes in frequency observed between the gaseous, liquid, and solid forms of water, for as the molecules become more closely packed together the fields will increase in strength ‡ The change in frequency observed in ice with decrease in temperature is also to be accounted for on these grounds, since the random motions at higher temperatures will always tend to weaken the influence of these internal fields and so allow the frequency to approach its unperturbed value as found in the gas

A résumé of most of the existing data on change in frequency with change of state in the Raman effect, as shown in Table II, is seen to give considerable support to this view. It will be noticed that all the molecules which exhibit

---

* 'Proc. Roy. Soc.,' A, vol. 125, p. 670 (1929).

† 'Phys. Rev.,' vol. 17, p. 20 (1921).

‡ The difference between the values obtained for the frequency in the gas as given by the Raman effect (3654 cm^{-1}.) and as given by infra-red measurements (3748 cm.$^{-1}$) can now be understood when it is recalled that the observations of the former were made at pressures of a few atmospheres, while those of the latter were made at pressures of a few millimetres of mercury.

Table II.—Change in Raman Frequency with Change in State.†

	Gas.	Liquid.	Solid.		Electric moment × 10^{18}
			Moderate temperature	Low temperature.	
H_2O	3650	3216 v.d. 3435 v d. 3582 v d.	3193 d. 3391 d 3549 d.	3090 s. 3135 s.	1.87
HCl	2886	2800 d			1.03
HBr	2558	2487 d			0.78
SO_2*	1152 1361 546	1146 s 1340 d. 526 d			1.6
H_2S	2615 d	2578 ?			
NH_3	3334	3216 s d. 3300 s.d. 3380 s d		3203 s d. 3369 s.	1.49
CO_2	1285.1 1387.7	1285.5 s 1387.5 s		1285 s. 1388 s.	0
CH_4	2914	2909 s		2906 s.	0
N_2O	1282 2226	1281.5 s. 2223.5 s.			0
H_2	4162	4149 s			0
O_2	1555	1550 s			0
N_2	2331	2326 s			0

v d, very diffuse, d, diffuse, s.d, slightly diffuse, s, sharp.

* The values for the frequencies at 1361 cm $^{-1}$ and 546 cm $^{-1}$ are deduced from the work of Bailey, Cassie and Angus ('Proc. Roy Soc.,' A, vol. 130, p 133 (1930)) on the infra-red absorption spectrum of the gas.

large frequency shifts in going from the gaseous to the liquid form also have large electric moments, while those with zero moment exhibit little or no change. This is exactly what one would expect, as it will be only in the case of strongly polar molecules that large, intermolecular fields can be developed. Next it will be seen that the change is greatest for these lines which are most diffuse, *e g.*, the water lines, the HCl and HBr lines, and the two SO_2 lines at

---

† The data in this table are taken from the work of various experimenters as collected in Kohlrausch's book "*Der Smekal-Raman Effekt*" (1931), or from the present paper, except for those on liquid and solid CO_2, which are to be found in a paper by McLennan and Smith, 'Canadian J. Res.,' *loc. cit.*

1361 cm.$^{-1}$ and 546 cm.$^{-1}$. It will be of interest to obtain the spectra of many of these polar molecules at −190° C., and the author hopes to undertake this work at an early date. When more accurate measurements of the widths of these lines are available it may be possible to make deductions regarding the magnitude of the intermolecular fields.

As regards the existence of associated molecules in water and ice, it would seem more satisfactory to assume that only two kinds of molecules* are present, viz., H_2O and $(H_2O)_2$ We accordingly attribute the line near 3200 cm $^{-1}$ to the $(H_2O)_2$ molecule, and look for an explanation of the other two frequencies in a doubling of the H_2O frequency arising from a condition of resonance degeneracy such as has been observed in CO_2. This is suggested by the fact that one of the other fundamental frequencies of water is known from infra-red and Raman spectra† to lie near 1650 cm.$^{-1}$ so that its first overtone may fall close enough to the fundamental for these two levels to interact. As the temperature is lowered, the position of the fundamental shifts considerably, as well as becoming more sharply defined, so that the interaction between the two levels may become negligible and only one frequency will be observed.

Methane.

The Raman spectrum of methane has been observed in the gaseous form,‡ and in the liquid form,§ but no observations seem to have been made on solid methane. Yet the Raman spectrum of solid methane is of particular interest in view of Pauling's prediction‖ that above 20° K. the molecules should rotate freely in the methane crystal. Although no pure rotation spectrum can be observed,¶ nevertheless it should be possible to pick up some of the rotation lines in the vibration-rotation band which has its centre at 3022 cm $^{-1}$ and was observed by Dickinson, Dillon, and Rassetti in the Raman spectrum of the gas.

---

* This is also in agreement with the conclusion of Kinsey and Sponslor ('Phys. Rev.,' vol. 40, p. 1035 (1932)), from a study of the X-ray data that the polymer in ice and water is dihydrol.

† Plyler and Sleator, 'Phys. Rev.,' vol. 37, p. 1493 (1931), Johnston and Walker, *ibid.*, vol. 39, p. 535 (1932).

‡ Dickinson, Dillon and Rassetti, 'Phys. Rev.,' vol. 34, p. 582 (1929); Bhagavantam, 'Ind. J. Phys.,' vol. 6, p. 595 (1932).

§ McLennan, Smith and Wilhelm, 'Trans. Roy. Soc. Canada,' vol. 23, p 247 (1929); Daure, 'Ann. Physique,' vol. 12, p. 375 (1929).

‖ 'Phys. Rev.,' vol. 36, p. 430 (1930).

¶ Houston, 'Phys. Rev.,' vol. 41, p. 263 (1932).

Several attempts were made to do this, but without success. The continuous background due to reflections and scattering from the amorphous mass was sufficiently intense to mask all lines except a very intense one at 2906 cm.$^{-1}$ which was scattered strongly by each of the group of mercury lines at 3650 A. as well as by the line at 4047 A. This line obviously corresponds to the intense line observed in the gas at 2914 cm.$^{-1}$, and in the liquid at 2909 cm $^{-1}$. Some method of making a clear crystal of solid methane will have to be developed before the existence of these faint lines can be established.

Ammonia.

Only one intense line was obtained from solid ammonia at −190° C.; it was scattered from the strong line at 3650 A as well as from the line at 4047 A., the measured frequency shift being 3369 cm.$^{-1}$ in each case. There were fainter companion lines at 3203 cm.$^{-1}$ and 1585 cm.$^{-1}$ scattered only by the line at 4047 A. The latter were not so sharply defined as the more intense frequency at 3369 cm.$^{-1}$. When one compares these results with those in Table III on gaseous and liquid ammonia it is obvious that some fundamental change occurs in the molecular structure of ammonia as it is gradually condensed into the solid state. If one considers first only the three lines near 3300 cm.$^{-1}$, the most remarkable thing is the disappearance of the middle frequency in the solid. The fact that this frequency appears by itself in the gaseous spectra and has been definitely proved to be a fundamental frequency

Table III.—Comparison of the Spectra of Gaseous Liquid and Solid Ammonia.

	Frequencies in cm.$^{-1}$.				
Gaseous ammonia*—					
Infra-red	933 966	1630		3336	
Raman	933·4 964 3			3334	
Liquid ammonia† (room temperature —40° C, *circa*)— Raman	1070 (0)		3216 (2)	3300 (4)	3380 (0)
	1070 (0)	1580 (0)	3210 (4)	3310 (4)	3380 (4)
Solid ammonia (at −190° C)— Raman		1585 (0)	3203 (1)		3369 (4)

* Robertson and Fox, 'Proc. Roy. Soc,' A, vol. 120, p. 207 (1928), Barker, 'Phys. Rev,' vol. 33, p. 684 (1929), Amaldi and Placzek,' 'Naturwiss,' vol 20, p. 521 (1932).

† Daure, *loc. cit.*; Bhagavantam, 'Ind J. Phys,' vol 5, p. 54 (1930)

of the NH_3 molecule* is fairly strong evidence that solid ammonia is not composed of simple NH_3 molecules. Bhagavantam (*loc. cit.*) has already suggested that the lines at 3216 cm.$^{-1}$ and 3380 cm.$^{-1}$ in the liquid are due to a more complex molecule such as $(NH_3)_2$, and the above results on the solid support such an interpretation.

The origin of the lines at 1070 cm.$^{-1}$ and 1580 cm.$^{-1}$ is more difficult to decide. Dadieu and Kohlrausch† have assigned them to the $(NH_3)_2$ molecule, but Bhagavantam (*loc. cit.*) prefers to associate them with the NH_3 molecule, since the intensity of the one at 1070 cm.$^{-1}$ is not appreciably increased at the lower temperature, and observations on the infra-red spectrum of the gas have proved that the frequencies 933 cm^{-1}, 966 cm^{-1}, and 1630 cm.$^{-1}$ are fundamental frequencies of the NH_3 molecule. It is difficult to reconcile this view with the fact that the 1580 cm.$^{-1}$ frequency is present in the solid, and to explain why one frequency should increase as the change is made from the gaseous to the liquid form while the other decreases. The fact that no Raman frequencies were found in the gas by Amaldi and Placzek (*loc. cit.*) which might correspond to these, even with very long exposures which brought out several new lines, is also against their interpretation as NH_3 frequencies On the whole the evidence would favour the explanation given by Dadieu and Kohlrausch.

Carbon Tetrachloride

An exposure was made of solid carbon tetrachloride at the temperature of liquid air. It was found to scatter very strongly and give the same lines as it does in the liquid form at room temperature The contour of the lines seemed sharper at the lower temperature, and the results, Table IV, show that the relative intensity of the lines is apparently slightly altered by the change in temperature. The interpretation of this change in intensity with temperature is not at all clear. The anti-Stokes lines were, of course, not observed at the lower temperature.

Table IV.—Raman Lines of Carbon Tetrachloride

Temperature.	Frequencies cm.$^{-1}$.					
25° C.	217 (8)	313 (9)	459 (10)	760 (3)	791 (3)	1537 (1)
−190° C.	215 (3)	314 (4)	458 (6)	758 (0)	788 (0)	

---

* Dennison and Hardy, 'Phys. Rev.,' vol. 39, p. 939 (1932).

† 'Naturwiss,' vol. 18, p. 154 (1930).

Liquid Air.

Some photographs were taken of the light scattered with no observation tube and only liquid air in the Dewar flask When the liquid air was fresh, then Raman lines of liquid oxygen and liquid nitrogen could be obtained in a few hours, but if the liquid air was a few days old, then only the oxygen lines were observed. The values obtained are given in Table V and are slightly lower than those recorded by McLennan and McLeod.* These values are the mean of several taken from different plates and scattered by two different mercury lines, viz , those at 4047 A. and 4358 A and are probably correct to 1 cm $^{-1}$.

Table V.—Raman Effect in Liquid Air.

Observer.	Oxygen.	Nitrogen.
McLennan and McLeod	1552 cm. $^{-1}$	2330 cm $^{-1}$
Sutherland	1550 cm.$^{-1}$	2326 cm $^{-1}$

Acknowledgments.

My best thanks are due to Professor T. M. Lowry, F R.S , under whose guidance the work was begun, and to Dr. C. P. Snow for his stimulating interest and encouragement. This research was carried out, partly in the Physical Chemistry Laboratory, Cambridge, and partly in the Physics Laboratory of the University of Michigan. In the latter place I would gratefully acknowledge the kindness shown me by various members of the staff , more particularly am I indebted to Professor Sawyer for his interest in the experimental side, and to Professor Dennison for much helpful discussion of the results. I wish also to express my indebtedness to the Carnegie Trust for a Scholarship, and to the Commonwealth Fund of New York for a Fellowship.

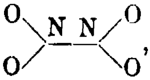
Summary.

(1) A new and very simple apparatus is described for the rapid examination of Raman spectra at very low temperatures.

(2) The complete Raman spectrum of solid nitrogen tetroxide has been obtained and its interpretation is discussed. The lines are divided into two classes ; those which are due to characteristic vibrations within a molecule

---

* 'Nature,' vol. 123, p. 160 (1929).

of the form , and those which are due to oscillations of this molecule about an equilibrium position in the crystal.

(3) An attempt to obtain the Raman spectrum of liquid ozone was unsuccessful. This negative result is shown to have important consequences in the determination of the form of the ozone molecule.

(4) The Raman spectrum of ice has been photographed at the temperature of liquid air, and found to differ markedly from that obtained at 0° C. It is shown that this can be accounted for by the existence of strong intermolecular fields in the crystal.

(5) This assumption of strong intermolecular fields is further developed and provides an explanation of the diffuseness of certain Raman lines which alter greatly in frequency between the gas, liquid, and solid phases.

(6) The evidence for the existence of associated molecules in water and in ice is discussed ; it is concluded that while there are strong reasons for believing that $(H_2O)_2$ molecules are present in water and in ice all the observed phenomena can be better explained without assuming the existence of higher aggregates.

(7) The spectrum of solid methane has been observed with a view to verifying Pauling's predictions on the rotation of the methane molecules in the crystal. The results are indecisive on this point.

(8) The Raman spectrum of solid ammonia has been obtained at the temperature of liquid air. The results indicate that the crystal is composed of some polymer of the simple NH_3 molecule such as $(NH_3)_2$.

(9) The results of observations on the Raman spectrum of carbon tetrachloride at the temperature of liquid air, and of observations on the Raman spectrum of liquid air are given. The former are found to agree (except for a slight change in relative intensity) with observations at room temperature, the latter to differ only very slightly from the results of earlier observers.

On the Electromagnetic Fields due to Variable Electric Charges and the Intensities of Spectrum Lines according to the Quantum Theory.

By V. Fock, University of Leningrad.

(Communicated by P. A. M Dirac, F.R S —Received March 18, 1933)

In a recent paper under the same title,† G. A. Schott finds that the usual method of calculating the intensities of spectrum lines is liable to lead to gross errors. Although one would expect from general reasons that any improvement of the usual method would only lead to small corrections and that, therefore, Schott's conclusion cannot be free from error, it may not be devoid of interest to show that a correct calculation *starting with Schott's formulæ* leads, in contradiction to Schott's conclusion, to the generally accepted classical result.

The aim of Schott's calculations is to obtain a quantum mechanical expression for the energy radiated by an atom, which is classically given by

$$\mathrm{R} = \frac{2}{3c^3}\left|\ddot{p}\left(t - \frac{r}{c}\right)\right|^2, \tag{1}$$

$p(t)$ being the electric moment of a dipole representing the atom. Schott considers the special case of a hydrogen atom emitting a radiation corresponding to a transition from the centrosymmetrical state

$$\psi_1 = \mathrm{A} f(r) \tag{2}$$

to the axiosymmetrical state

$$\psi_2 = \mathrm{B}\,(\mathbf{r}\mathbf{p}_1)\,g\,(r), \tag{3}$$

where

$$f(r) = e^{-\frac{r}{al}}\,\mathrm{L}'_l\left(\frac{2r}{al}\right); \qquad g\,(r) = e^{-\frac{r}{ak}}\,\mathrm{L}'''_{k+1}\left(\frac{2r}{ak}\right) \tag{4}$$

are wave functions expressed in terms of derivatives of Laguerre's polynomials L_m and corresponding to principal quantum numbers $n_1 = l$ and $n_2 = k$. The letter a denotes Bohr's hydrogen radius. We adopt here Schott's notations and put according to him the velocity of light c equal to unity.

Schott finds by his method that the energy radiated is equal to

$$\mathrm{R} = \tfrac{2}{3}\mathrm{I}^2, \tag{5}$$

---

† 'Proc. Roy. Soc.,' A, vol. 139, p. 37 (1933).

where

$$I = -\frac{8\pi e A B \nu a k^2 l}{l^2 - k^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{i\mu(t-r-\tau)}\, i \sin(2\pi\nu\tau + \varepsilon) \sqrt{\frac{2\pi}{\mu}}\, Q\, d\tau\, d\mu \tag{6}$$

and

$$Q = \int_0^{\infty} e^{-\alpha s} \left\{ L'_l\left(\frac{2s}{al}\right) - 2L''_l\left(\frac{2s}{al}\right) \right\} L'''_{k+1}\left(\frac{2s}{ak}\right) J_{3/2}(\mu s)\, s^{3/2}\, ds, \tag{7}$$

α denoting the quantity

$$\alpha = \frac{1}{al} + \frac{1}{ak} \tag{8}$$

and $J_{3/2}$ denoting Bessel's function of order 3/2.

On the other hand, Schott gives the following expression for the electric moment p:

$$p = \tfrac{2}{3}\pi e A B \cos(2\pi\nu\tau + \varepsilon) \,.\, P, \tag{9}$$

where

$$P = \int_0^{\infty} e^{-\alpha s} L'_l\left(\frac{2s}{al}\right) L'''_{k+1}\left(\frac{2s}{ak}\right) s^4\, ds. \tag{10}$$

Equations (6), (7), (9), and (10) are taken without any change from Schott's paper (his formulæ (24), (25), (26), and (27)).

We shall now calculate the quantity I and express it in terms of p.

Noting that

$$\lim_{\mu \to 0} \frac{3}{2\mu} \sqrt{\frac{2\pi}{\mu}}\, J_{3/2}(\mu s) = s^{3/2}$$

is a finite quantity, we introduce the function

$$Q^*(\mu) = \frac{3}{2\mu} \sqrt{\frac{2\pi}{\mu}}\, Q, \tag{11}$$

where Q is defined by (7). The quantity $Q^*(0)$ is finite and equal to

$$Q^*(0) = \int_0^{\infty} e^{-\alpha s} \left\{ L'_l\left(\frac{2s}{al}\right) - 2L''_l\left(\frac{2s}{al}\right) \right\} L'''_{k+1}\left(\frac{2s}{ak}\right) s^3\, ds, \tag{12}$$

or, expressed in terms of f and g, equations (4),

$$Q^*(0) = -al \int_0^{\infty} f'(r)\, g(r)\, r^3\, dr. \tag{13}$$

We have also, by (10) and (4)

$$P = \int_0^{\infty} f(r)\, g(r)\, r^4\, dr. \tag{14}$$

Now, f and g satisfy the equations

$$f''(r) + \frac{2}{r} f'(r) = \left(\frac{1}{a^2 l^2} - \frac{2}{ar}\right) f(r),$$

$$g''(r) + \frac{4}{r} g'(r) = \left(\frac{1}{a^2 k^2} - \frac{2}{ar}\right) g(r).$$

It follows that

$$\frac{d}{dr}\left[r^4 (f'g - g'f)\right] = 2r^3 f'g - \frac{l^2 - k^2}{a^2 l^2 k^2} r^4 fg.$$

Consequently

$$\int_0^\infty fgr^4\,dr = \frac{2k^2 l^2 a^2}{l^2 - k^2} \int_0^\infty f'gr^3\,dr, \tag{15}$$

or, by (13) and (14),

$$P = -\frac{2k^2 la}{l^2 - k^2} Q^*(0). \tag{16}$$

This transformation is a consequence of the general quantum-mechanical equations of motion.

Let us now consider the expression (6) for I. Introducing in place of τ the new variable of integration

$$u = \tau - t + r$$

and remembering that

$$\sqrt{\frac{2\pi}{\mu}}\, Q = \tfrac{2}{3} \mu Q^*(\mu)$$

is an odd function of μ, we may write

$$I = -\frac{8\pi e A B \nu a k^2 l}{l^2 - k^2} \cos(2\pi\nu(t - r) + \varepsilon)$$
$$\times \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \sin \mu u \sin 2\pi\nu u \, \tfrac{2}{3}\mu Q^*(\mu)\, du\, d\mu. \tag{17}$$

But the double integral is the Fourier expansion of the function

$$\tfrac{8}{3}\pi^2 \nu Q^*(2\pi\nu).$$

Consequently, equation (17) may be written

$$I = -\frac{64}{3}\pi^3 \nu^2 \frac{e A B a k^2 l}{l^2 - k^2} \cos(2\pi\nu(t - r) + \varepsilon)\, Q^*(2\pi\nu), \tag{18}$$

or, introducing the expression (9) for $p(\tau)$ and using (16),

$$I = 4\pi^2 \nu^2 p(t - r) \frac{Q^*(2\pi\nu)}{Q^*(0)}, \tag{19}$$

or finally

$$I = -p(t-r)\frac{Q^*(2\pi\nu)}{Q^*(0)}. \tag{20}$$

Now $2\pi\nu$ is a small quantity of the order of the fine structure constant

$$\alpha = \frac{2\pi e^2}{ch} = \frac{1}{137,3}$$

Since $Q^*(2\pi\nu)$ is an even function of $2\pi\nu$, it differs from $Q^*(0)$ by quantities of the order α^2 which are to be neglected in Schrodinger's theory. (These quantities are also neglected in Schott's calculations.) Consequently we have the simple result

$$I = -p(t-r). \tag{21}$$

Introducing this in (5) and remembering that, following Schott, we have put $c = 1$, we see that equation (5) obtained by Schott's method exactly coincides with the generally accepted classical expression (1).

Schott's criticism of the usual procedure is therefore based on a misunderstanding. As to the method used by Schott, it is not new, since it was used in 1927 by Klein,[†] the only novelty is the introduction of double Fourier expansions which is not necessary in our case since we are dealing with simple harmonic functions of the time.

Summary.

The paper is a criticism of a paper by Schott.[‡] It is shown that Schott's conclusions cannot be substantiated, and that the usual methods of calculating the intensities of spectral lines are correct.

[†] 'Z. Physik,' vol. 41, p. 407 (1927).

[‡] 'Proc. Roy. Soc.,' A, vol. 139, p. 37 (1933).

The Coefficients of Absorption and Opacity of a Partially Degenerate Gas.

By Bertha Swirles, Ph.D., Imperial College of Science and Technology.

(Communicated by E. A. Milne, F.R.S.—Received March 18, 1933.)

§ 1. The absorption of radiation in a highly degenerate gas has been investigated by Chandrasekhar,* Majumdar,† and the present writer.‡ Chandrasekhar§ has also considered very fully the absorption in a classical gas. It appears that the absorption in a highly degenerate gas is much less than that which would be calculated for a classical gas at the same density and temperature. The present paper is concerned with the absorption in a gas for which the degeneracy, as measured by Sommerfeld's parameter, is small. As in a stellar configuration with a degenerate-gas core there must be a transition zone, where the gas is only slightly degenerate, it may be of value for astrophysical purposes to know the behaviour of the absorption coefficient for incipient degeneracy.

In § 2 the coefficients of absorption and opacity due to "free-free" transitions are evaluated; in §§ 3 and 4 the contributions to the atomic absorption coefficient due to "bound-free" transitions from the lower energy levels are calculated. In § 5 the absorption coefficient per normal hydrogen atom is found and the mass absorption coefficient deduced, and in § 6 the extrapolation for other atoms is discussed.

It has not been found possible to make a complete evaluation of the opacity, owing to a difficulty to be explained later.

The gas is supposed to consist of positively charged nuclei of charge Ze and electrons. The state of ionization is given by a modification of the standard ionization formula. The electrons alone are supposed to be degenerate. This is justifiable owing to the greater mass of the nuclei. The probability that a cell of energy E is empty is then given by

$$\frac{e^{\alpha+\mathrm{E}/k\mathrm{T}}}{e^{\alpha+\mathrm{E}/k\mathrm{T}}+1}, \tag{1}$$

---

* 'Proc. Roy. Soc.,' A, vol. 133, p. 241 (1931). Referred to as C I.

† 'Astr. Nachr.,' vol. 243, p. 6 (1931); vol. 247, p. 12 (1932); Kothari and Majumdar, 'Astr. Nachr.,' vol. 244, p. 66 (1931).

‡ 'Mon. Not. R. Astr. Soc.,' vol. 91, p. 857 (1931).

§ 'Proc. Roy. Soc.,' A, vol. 135, p. 472 (1932). Referred to as C II.

where, if N_e is the number of electrons per unit volume and T the temperature, α is given in terms of Sommerfeld's degeneracy parameter S by the equation

$$S = \frac{h^3N_e}{2\,(2\pi mkT)^{3/2}} = \frac{2}{\pi^{1/2}}\int_0^{\infty}\frac{x^{1/2}\,dx}{e^{\alpha+x}+1}. \tag{2}$$

§ 2. *The Coefficients of Absorption and Opacity due to "free-free" Transitions.*—Using the results of Gaunt* the rate of absorption of energy from radiation of frequency ν by electrons in states of positive energy in a range (E, E + dE) in the field of a nucleus with charge Ze, is, per unit volume

$$a(\nu)_{\mathrm{E}} = \frac{32\pi^2Z^2e^6}{3\sqrt{3}hc\mathrm{E}^{1/2}\nu^3}\cdot g\cdot\frac{e^{\alpha+(\mathrm{E}+h\nu)/k\mathrm{T}}}{(e^{\alpha+\mathrm{E}/k\mathrm{T}}+1)\,(e^{\alpha+(\mathrm{E}+h\nu)/k\mathrm{T}}+1)}, \tag{3}$$

where $g \to 1$ when E, E′ (= E + $h\nu$) $\to 0$. It appears also that except for very high frequencies $g = 1$ gives a good approximation when E alone tends to zero. For extreme degeneracy, an approximation to the repartition of the electrons among the phase cells is given in the following way. For energies less than the "threshold energy," $\frac{h^2}{8m}\left(\frac{3N_e}{\pi}\right)^{2/3}$, all phase cells are occupied, while for energies exceeding this value all phase cells are vacant. Chandrasekhar (C I) points out that the maximum of the Planck curve for the radiation then falls well within the "threshold energy" and infers that the greater part of the contribution to the opacity is from transitions such that E > $h\nu$. He therefore uses a formula due to Gaunt suitable for E ≫ $h\nu$. This argument cannot be applied with the same force to the incipient degeneracy considered here. The maximum of the weighting function used in evaluating the opacity (Rosseland mean) lies near the frequency $4k\mathrm{T}/h$, and that for the "straight mean" near $2\cdot8\,k\mathrm{T}/h$.† The "threshold energy" is given approximately by $\frac{h^2}{8m}\left(\frac{3N_e}{\pi}\right)^{2/3}$ and the condition for both these maxima to be *outside* the threshold energy becomes

$$2\cdot3 > \mathrm{S}^{2/3} \quad\text{or}\quad \mathrm{S} < 3\cdot5. \tag{4}$$

We shall be mainly concerned with values of S which satisfy this condition, and for values of S outside this range it is not clear that the Gaunt formula for E ≫ $h\nu$ is more suitable than (3). Using therefore (3) with $g = 1$, the atomic absorption coefficient due to "free-free" transitions is, for frequency ν, equal to

---

* 'Phil. Trans.,' A, vol. 229, p. 163 (1930).

† *Cf.* Majumdar, *loc. cit.* (1932).

$$a_1(\nu) = \frac{A}{\nu^3\left(1 - e^{-h\nu/kT}\right)}\left\{h\nu + kT\log\left(\frac{e^{\alpha}+1}{e^{\alpha+h\nu/kT}+1}\right)\right\}, \tag{5}$$

where

$$A = \frac{32\pi^2 Z^2 e^6}{3\sqrt{3}h^4 c}.^{*} \tag{6}$$

The "straight mean" is therefore given by

$$a_0 = \frac{A\displaystyle\int_0^\infty \frac{\left\{h\nu + kT\log\left(\dfrac{e^{\alpha}+1}{e^{\alpha+h\nu/kT}+1}\right)\right\}}{e^{h\nu/kT}-1}\,d\nu}{\displaystyle\int_0^\infty \frac{\nu^3\,d\nu}{e^{h\nu/kT}-1}} \tag{7}$$

$$= \frac{160 Z^2 e^6}{\pi^2\sqrt{3}hc\,(kT)^2}\int_0^\infty \frac{x + \log\left(\dfrac{e^{\alpha}+1}{e^{\alpha+x}+1}\right)}{e^x-1}\,dx. \tag{8}$$

The integral in (8) is equal to

$$\log_e 10\int_0^\infty \frac{\log_{10}\left(1+e^{-\alpha}\right) - \log_{10}\left(1+e^{-(\alpha+x)}\right)}{e^x-1}\,dx = \vartheta, \text{ say.} \tag{9}$$

To evaluate ϑ for $\alpha = 0{\cdot}5$, 0, $-0{\cdot}5$, -1 and -5, the values of the integrand were tabulated at intervals of 0·5 in x and the integration made with a planimeter. The results are given in Table I, columns (5) and (7)

The atomic opacity coefficient or Rosseland mean is defined as a_1 where

$$a_1^{-1} = \frac{\displaystyle\int_0^\infty \frac{1}{a(\nu)\left(1-e^{-h\nu/kT}\right)}\frac{\partial I_\nu}{\partial T}\,d\nu}{\displaystyle\int_0^\infty \frac{\partial I_\nu}{\partial T}\,d\nu} \tag{10}$$

where $a(\nu)$ is the rate of absorption of energy from radiation of frequency ν due to all transitions, "bound-free" as well as "free-free." The opacity is not therefore an additive function† of the contributions from all transitions and its complete evaluation presents great difficulties.

---

* In my previous paper a factor 2 has been omitted here in error. In that paper the formula (29) should be multiplied by 2, but (30) remains unaltered.

† *Cf* Milne, 'Mon. Not. R. Astr. Soc.,' vol. 85, p. 979 (1925).

We shall obtain here only the coefficient of opacity due to "free-free" transitions in order to compare it with the corresponding absorption coefficient. We obtain

$$a_1 = \frac{128\pi^6 Z^2 e^6}{45\sqrt{3}hc\,(kT)^2}\left|\int_0^\infty \frac{x^7 e^{-x}}{(1-e^{-x})^2\left\{x+\log\left(\frac{e^{\alpha}+1}{e^{\alpha+x}+1}\right)\right\}}\,dx\right|^{-1}. \tag{11}$$

For convenience in computation, the integral in the denominator of (37) may be written*

$$\frac{1}{\log(1+e^{-\alpha})}\left[\int_0^\infty \frac{x^7 e^{-x}}{(1-e^{-x})^2}\,dx\right.$$
$$\left.+\frac{1}{2}\int_0^\infty \frac{x^7\log_{10}(1+e^{-(\alpha+x)})}{(\cosh x-1)\{\log_{10}(1+e^{-\alpha})-\log_{10}(1+e^{-(\alpha+x)})\}}\,dx\right] \tag{12}$$
$$=\frac{1}{\log(1+e^{-\alpha})}\left[5082+\frac{1}{2}\int_0^\infty f(x,\alpha)\,dx\right]. \tag{13}$$

For $\alpha = 0{\cdot}5,\ 0,\ -0{\cdot}5,\ -1,\ -5$, $f(x, \alpha)$ was tabulated from $x = 0$ to $x = 10$ at intervals of $0{\cdot}5$. The integration was then made by means of a planimeter. The error is estimated at about 2% As for all values of α considered $\frac{1}{2}\int_0^\infty f(x,\alpha)\,dx$ is much less than 5082 this degree of accuracy is sufficient The results are given in Table I, column 4 In column 2, Table I, the values of $S\ (= h^3N/\{2\,(2\pi mkT)^{\frac{3}{2}}\})$, corresponding to the required values of α, are given.

$$S = \frac{2}{\pi^{\frac{1}{2}}}\int_0^\infty \frac{x^{\frac{1}{2}}\,dx}{e^{\alpha+x}+1}. \tag{14}$$

For $\alpha = 0,\ 0{\cdot}5$, this is equal to

$$\frac{e^{-\alpha}}{1^{\frac{3}{2}}}-\frac{e^{-2\alpha}}{2^{\frac{3}{2}}}+\frac{e^{-3\alpha}}{3^{\frac{3}{2}}}-\ \ldots . \tag{15}$$

For $\alpha = -0{\cdot}5,\ -1$, the integrand was tabulated from $x = 1$ to $x = 10$ at intervals of $0{\cdot}5$ in x and the integration made with a planimeter For

$$\frac{2}{\pi^{\frac{1}{2}}}\int_0^\infty \frac{x^{\frac{1}{2}}\,dx}{e^{-5+x}+1} = \frac{5^{\frac{3}{2}}}{\Gamma(\frac{5}{2})}+2\cdot 5^{\frac{3}{2}}\left[\frac{\overline{12}}{5^2\Gamma(\frac{1}{2})}+\frac{\overline{720}}{5^4\Gamma(-\frac{1}{2})}+\frac{\overline{30240}}{5^6\Gamma(-\frac{3}{2})}\right] \tag{16}$$

with an error of the order of $e^{-5} = 0{\cdot}00674$.† The sum of the series on the right-hand side of (16) is 8·816 and this may be taken as the value of the integral with an error of less than 1 %.

---

* I wish to thank Professor D R. Hartree, F R.S., for discussion with regard to the evaluation of this integral.

† See Brillouin, "Statistiques Quantiques," vol. 2, p. 387.

The values of a_1 multiplied by $hc\,(k\mathrm{T})^2/\mathrm{Z}^2e^6$ are given in Table I, column 6. The values of a_1 and a_0 for very high degeneracy ($\alpha \to -\infty$) are included in Table I, using the formulæ (see C I)

$$ {}_{\mathrm{D}}a_1 = \frac{56}{15\sqrt{3}} \cdot \frac{\mathrm{Z}^2e^6}{ch\,(k\mathrm{T})^2} \tag{17} $$

$$ {}_{\mathrm{D}}a_0 = \frac{80}{3\sqrt{3}} \cdot \frac{\mathrm{Z}^2e^6}{ch\,(k\mathrm{T})^2} \tag{18} $$

It is seen that already for $\alpha = -5$, the values are very little different from those for high degeneracy.

For a classical gas ($\alpha \to +\infty$), (C I)

$$ {}_{c}a_1 = 0{\cdot}155\,\frac{\mathrm{Z}^2e^6h^2}{c\,(2\pi m)^{3/2}} \cdot \frac{\mathrm{N}_e}{(k\mathrm{T})^{7/2}} \tag{19} $$

$$ {}_{c}a_0 = 4{\cdot}68\,\frac{\mathrm{Z}^2e^6h^2}{c\,(2\pi m)^{3/2}} \cdot \frac{\mathrm{N}_e}{(k\mathrm{T})^{7/2}}\,. \tag{20} $$

Now since $\mathrm{S} = h^3\mathrm{N}_e/[2\,(2\pi mk\mathrm{T})^{3/2}]$, the values of a_1 and a_0 are unaltered when multiplied by $\dfrac{1}{\mathrm{S}}\,\dfrac{h^3\mathrm{N}_e}{2\,(2\pi mk\mathrm{T})^{3/2}}$. By multiplying by this factor we obtain the

Table I

α.	S.	$\log(1+e^{-\alpha})$	$5082+\frac{1}{2}\int_0^\infty f(x,\alpha)\,dx$	ϑ.	$a_1\,\dfrac{hc\,(k\mathrm{T})^2}{\mathrm{Z}^2e^6}$.	$a_1\,\dfrac{hc\,(k\mathrm{T})^2}{\mathrm{Z}^2e^6}$
0 5	0 508	0 4740	5110 8	0 433	0 146	4 52
0	0 767	0 6931	5114 5	0 610	0 214	5 71
−0 5	1 117	0 9742	5120 0	0 797	0 300	7 46
−1	1 568	1 3132	5128 3	0 983	0 404	9 20
−5	8 816	5 0066	5482 0	1 609	1 44	15 06
−∞	∞	∞	∞	1 645	2 15	15 39

Table II.

α	S	$a_1\,\dfrac{c\,(2\pi m)^{3/2}\,(k\mathrm{T})^{7/2}}{\mathrm{Z}^2e^6\,\mathrm{N}}$	$a_1/a_{1(\infty)}$	$a_0\,\dfrac{c\,(2\pi m)^{3/2}\,(k\mathrm{T})^{7/2}}{\mathrm{Z}^2e^6\mathrm{N}}$.	$a_0/a_{0(\infty)}$.
∞	0	0 155	1	4 68	1
0 5	0 508	0 144	0 93	4 45	0·96
0	0 767	0 139	0 90	3 72	0·79
−0 5	1 117	0 134	0 87	3·34	0·71
−1	1·568	0 129	0 83	2·93	0 63
−5	8 816	0 082	0 53	0 85	0·1

expressions for a_1 and a_0 in a form suitable for comparison with the classical expressions. In Table II, columns 3 and 5, the values of a_1 and a_0 multiplied by $c\,(2\pi m)^{3/2}\,(k\mathrm{T})^{7/2}/(\mathrm{Z}^2e^6h^2\mathrm{N}_e)$, and in columns 4 and 6 the ratios of a_1 and a_0 to the values which would be computed for them from the formulæ (19) and (20) for a classical gas, are given.

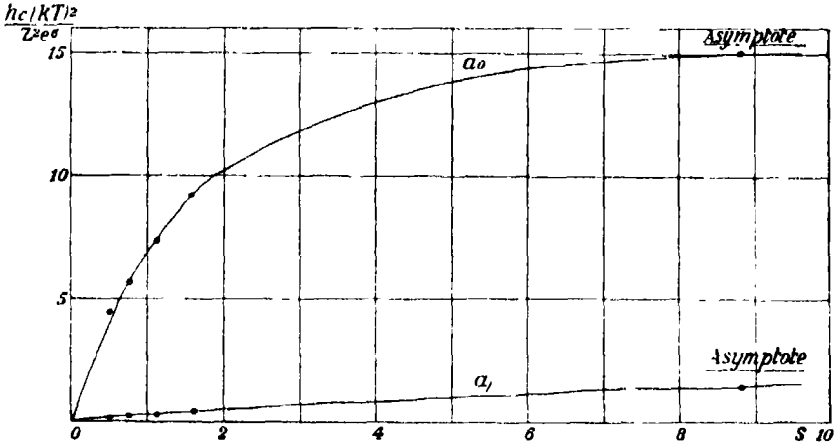


FIG. 1.

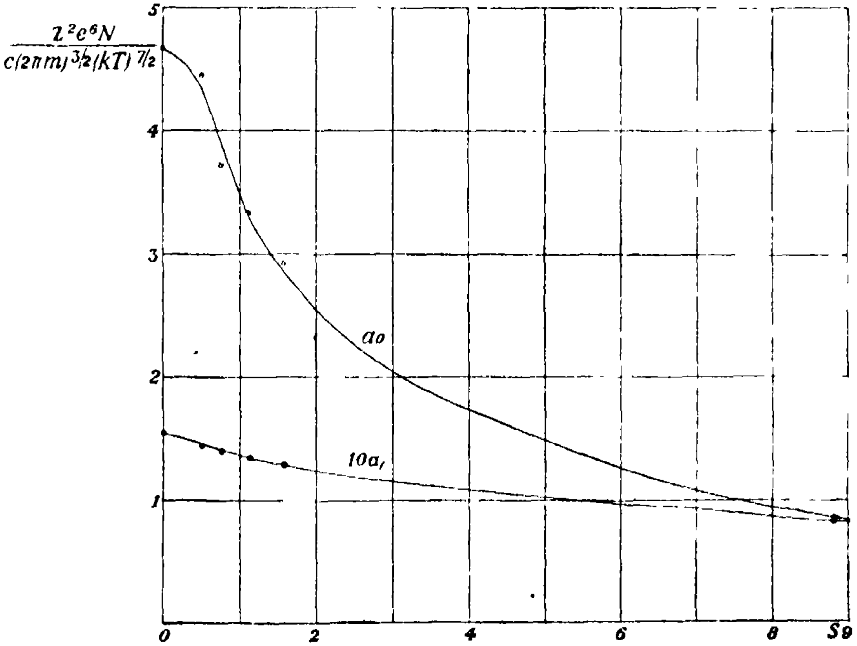


FIG. 2.

These results are also shown in figs. 1 and 2.

§ 3. *Absorption Coefficient due to the two K-electrons.*—Using the result of Nishina and Rabi,* and Stobbe,† the rate of absorption of radiation of frequency ν by the two K-electrons in the field of a nucleus of charge Z is found to be

$$(a_\nu)_k = \frac{2^8\pi e^2}{3cm}\cdot\frac{\nu_1{}^3}{\nu^4}\cdot\frac{\exp\,[-4\{\nu_1/(\nu-\nu_1)\}^{\frac12}\tan^{-1}\{(\nu-\nu_1)/\nu_1\}^{\frac12}]}{1-\exp.\,[-2\pi\{\nu_1/(\nu-\nu_1)\}^{\frac12}]}\;\frac{\exp.\,[\alpha+h\,(\nu-\nu_1)/k\mathrm{T}]}{\exp\,[\alpha+h\,(\nu-\nu_1)/k\mathrm{T}]+1}. \qquad (21)$$

where

$$\nu_n = \frac{\mathrm{R}Z^2}{n^2}, \qquad \mathrm{R} = \frac{2\pi^2 me^4}{h^3}.$$

Therefore the atomic absorption coefficient (the "straight mean") is given by

$$a_k = \frac{\displaystyle\int_{\nu_1}^{\infty}(a\nu)_\kappa\,\nu^3 e^{-h\nu/k\mathrm{T}}\,d\nu}{\displaystyle\int_0^{\infty}\frac{\nu^3}{e^{h\nu/k\mathrm{T}}-1}\,d\nu} \qquad (22)$$

Putting

$$\{(\nu-\nu_1)/\nu_1\}^{\frac12} = x \quad\text{and}\quad \frac{h\nu_1}{k\mathrm{T}} = y_1,$$

(22) reduces to

$$a_k = \frac{2^9\,.\,5}{\pi^3}\cdot\frac{e^2}{cm\nu_1}\;y_1^{4_{\prime}\,-\,\prime\prime_1}\int_0^{\infty}\frac{x\exp.\,[-(4/x)\tan^{-1}x]\,dx}{(x^2+1)\,[1-\exp.\,(-2\pi/x)]\,[\exp.\,(y_1x^2)+\exp.\,(-\alpha)]}. \qquad (23)$$

This reduces to equation (17) (C II) when e^α tends to $+\infty$ and may be evaluated in a similar way under certain conditions. The range of integration is divided into equal intervals $(k, k+1)$ where k is an integer, and in each of these intervals the expression

$$f(x) = \frac{x}{x^2+1}\cdot\frac{\exp\,[-(4/x)\tan^{-1}x]}{1-\exp.\,(-2\pi/x)}. \qquad (24)$$

is represented by a linear approximation

$$m_k x + c_k,$$

so that, to the extent to which this approximation is valid,

$$\int_k^{k+1}\frac{f(x)\,dx}{e^{y_1x^2}+e^{-\alpha}} \sim \frac{f(k+1)-f(k)}{2y_1}\Big[e^{-\alpha}\log\,(1+e^{-\alpha-y_1x^2})\Big]_k^{k+1}$$
$$+\,[(k+1)f(k)-kf(k+1)]\int_k^{k+1}\frac{dx}{e^{y_1x^2}+e^{-\alpha}}. \qquad (25)$$

---

* 'Verh. deuts. phys. Ges.,' vol. 9, p. 8 (1928).

† 'Ann. Physik,' vol. 7, p. 661 (1930).

The coefficients m_k, c_k are the same as in (18) (C II) and I am indebted to Mr. Chandrasekhar for checking his calculation of them. Then,

$$\begin{aligned}\int_0^\infty \frac{f(x)\,dx}{e^{y_1x^2}+e^{-\alpha}} = \frac{e^\alpha}{y_1}\{&0\cdot01082\,[\log(1+e^{-\alpha})-\log(1+e^{-\alpha-y_1})]\\ &+0\cdot01201\,[\log(1+e^{-\alpha-y_1})-\log(1+e^{-\alpha-4y_1})]\\ &+0\cdot00953\,[\log(1+e^{-\alpha-4y_1})-\log(1+e^{-\alpha-9y_1})]\\ &+0\cdot00710\,[\log(1+e^{-\alpha-9y_1})-\log(1+e^{-\alpha-16y_1})]\\ &+\\ &-\Big\{0\cdot00237\int_1^2\frac{dx}{e^{y_1x^2}+e^{-\alpha}}-0\cdot00753\int_2^3\frac{dx}{e^{y_1x^2}+e^{-\alpha}}\\ &\qquad-0\cdot02211\int_3^4\frac{dx}{e^{y_1x^2}+e^{-\alpha}}+\ldots\Big\}.\end{aligned} \tag{26}$$

We must now consider just what range of vaules for α and y_1 we require. Ignoring the effect of the finite volumes occupied by the atoms in their discrete quantum states the ionization formula for a gas can be obtained in the form

$$\frac{n_{r+1}}{n_r} = \frac{q_{r+1}}{q_r}\,e^{\alpha-\psi_r/kT},\text{*} \tag{27}$$

where n_r is the number of atoms $(r-1)$ times ionized, q_r is the statistical weight and ψ_r is the rth ionization potential. From this it would appear that as $\alpha \to -\infty$ (high degeneracy) the gas cannot be ionized But in order to account for the very high densities in white dwarf stars it is necessary to assume that the gas is very highly ionized. For such high densities (27) will no longer be valid since it becomes necessary to take into account the volumes occupied by the atoms.

It has been suggested by various writers† that for very high densities the ionization is determined by the fact that the volume which can be occupied by an ion is restricted. When the density is such that

$$\begin{pmatrix}\text{Factor of order}\\ \text{unity}\end{pmatrix} \times \begin{pmatrix}\text{Number of nuclei}\\ \text{per unit volume}\end{pmatrix} \times \begin{pmatrix}\text{Volume of } (r-1)\\ \text{times ionized atom}\end{pmatrix} > 1, \tag{28}$$

then the atoms must on an average be more than $(r-1)$ times ionized.

We shall now put the condition (28) in another form. Let an $(r-1)$

---

* *Cf.* Milne, 'Mon. Not. R. Astr Soc.,' vol. 90, p. 769 (1930).

† Kothari and Majumdar, 'Astr. Nachr.,' vol. 244, p. 66 (1931) where other references are given.

times ionized atom have a series electron describing a non-penetrating orbit in the field of a core with charge $Z_{\rm eff}$. Then the orbit corresponding to the ground state has radius approximately $a_{\rm H}/Z_{\rm eff}$, where $a_{\rm H}$ is the radius of the first hydrogen orbit. If ψ_r be the rth ionization potential

$$\psi_r = \frac{1}{2} \frac{Z_{\rm eff}{}^2 e^2}{a_{\rm H}}.$$

If now we consider each ion to occupy a sphere of radius $\kappa a_{\rm H}/Z_{\rm eff}$, where $\kappa > 1$, the condition (28) becomes, if N is the number of nuclei per unit volume,

$$\mathrm{N} \,.\, \tfrac{4}{3}\pi \left(\frac{\kappa a_{\rm H}}{Z_{\rm eff}}\right)^3 > 1 \qquad (29)$$

If N is the number of free electrons,

$$\mathrm{N}_e \sim (r-1)\,\mathrm{N}$$

and

$$k\mathrm{T}\mathrm{S}^{\frac{2}{3}} = \frac{h^2}{2\pi m}\left(\frac{\mathrm{N}_e}{2}\right)^{\frac{2}{3}}.$$

Then since $a_{\rm H} = h^2/4\pi^2 m e^2$, (29) becomes

$$\frac{\kappa^2}{4\pi}\left(\frac{8\pi}{3\,(r-1)}\right)^{\frac{2}{3}} \mathrm{S}^{\frac{2}{3}} > \psi_r/k\mathrm{T}. \qquad (30)$$

This is the condition for pressure ionization to occur. The numerical factor is less than unity and we may infer that pressure ionization will not occur, *i.e.*, we may apply (27) provided that

$$\psi_r/k\mathrm{T} > \mathrm{S}^{\frac{2}{3}}.$$

We shall therefore impose the condition

$$y_1 > \mathrm{S}^{\frac{2}{3}}. \qquad (31)$$

The range of values of S which we particularly wish to investigate is that between 0 and 3; calculations on the "standard model"* of a star with degenerate-gas core, give a value $\mathrm{S} = 2{\cdot}97$ at the interface.†

We find that for $\alpha > -1$ (corresponding to $\mathrm{S} < 1{\cdot}6$) and $y_1 > 3$ (thus satisfying (31)) a good approximation to (26) is given by the first term

$$0{\cdot}01082 \frac{e^{\alpha}}{y_1} \log\left(1 + e^{-\alpha}\right).$$

---

* Milne, 'Mon Not. R. Astr. Soc.,' vol. 91, p. 1 (1930).

† *Cf.* Kothari, 'Mon. Not. R. Astr. Soc.,' vol. 93, p. 71 (footnote) (1932).

For larger values of S the values of y_1 required by (31) will be such as to correspond to extremely low temperatures.

The contribution to the absorption coefficient by *one* K-electron is

$$\frac{2^8 . 5}{\pi^3} \cdot \frac{e^2}{cm\nu_1} \cdot y_1^3 e^{-y_1} \times 0{\cdot}01082 e^{\alpha} \log (1 + e^{-\alpha}),$$

whilst in a classical gas it is

$$\frac{2^8 . 5}{\pi^3} \cdot \frac{e^2}{cm\nu_1} \cdot y_1^3 e^{-y_1} \times 0{\cdot}01082.$$

§ 4. *The Coefficient of Absorption due to the Eight* L-*electrons.*—The rate of absorption of radiation of frequency ν by the eight L-electrons in the field of a nucleus of charge Z is given by

$$(a_\nu)_{\mathrm{L}} = \frac{2^{11}\pi e^2}{3cm} \frac{\nu_2^3}{\nu^4} \left(1 + 6 \frac{\nu_2}{\nu} + 8 \frac{\nu_2^2}{\nu^2}\right)$$

$$\times \frac{\exp. [-8\{\nu_2/(\nu - \nu_2)\}^{\frac{1}{2}} \tan^{-1}\{(\nu - \nu_2)/\nu_2\}^{\frac{1}{2}}]}{1 - \exp. [-4\pi\{\nu_2/(\nu - \nu_2)\}^{\frac{1}{2}}]} \cdot \frac{\exp. [\alpha + h(\nu - \nu_2)/k\mathrm{T}]}{\exp. [\alpha + h(\nu - \nu_2)/k\mathrm{T}] + 1}. \tag{35}$$

Putting

$$x = \{(\nu - \nu_2)/\nu_2\}^{\frac{1}{2}} \quad \text{and} \quad \frac{h\nu_2}{k\mathrm{T}} = y_2$$

the atomic absorption coefficient is given by

$$a_{\mathrm{L}} = \frac{2^{12} . 5}{\pi^3} \cdot \frac{e^2}{cm\nu_2} \cdot y_2^4 e^{-y_2} \cdot \frac{x}{x^2 + 1} \left\{1 + \frac{6}{x^2 + 1} + \frac{8}{(x^2 + 1)^2}\right\}$$

$$\times \frac{}{(1 - e^{-4\pi/x})(e^{y_2 x^2} + e^{-\alpha})} dx. \tag{36}$$

By similar considerations to those employed for the K-electrons it is seen that the values required are

$$y_2 \sim 3, \qquad \alpha > -1,$$

and by the same argument as before a good approximation to the integral in (36) is given by

$$\frac{0{\cdot}00280}{y_2} e^{\alpha} \log (1 + e^{-\alpha}). \tag{37}$$

As in the case of the K-electrons, the contribution to the absorption coefficient by an L-electron is $e^{\alpha}\log(1+e^{-\alpha})$ times the value given by the classical formula.

---

* Mr. Chandrasekhar has kindly sent me the corrected form of the result given by Stobbe, *loc. cit.*

§ 5. *The Absorption Coefficient due to Hydrogen Atoms.*—Let

n = total number of hydrogen atoms per cm.^{3},

n_l = number per cm.^{3} in quantum state l,

N_ι = number of hydrogen ions per cm.^{3},

N_e = total number of free electrons per cm.^{3},

a = coefficient of absorption per normal hydrogen atom,

The contribution to a due to the "free-free" transitions is

$$\frac{\mathrm{N}_\iota}{n_1}\,\,. \tag{38}$$

If the electrons alone are degenerate

$$\frac{\mathrm{N}_\iota}{n_1} = \tfrac{1}{2}e^{-(\alpha_\iota-\alpha_1)}, \tag{39}$$

where

$$\alpha_\iota - \alpha_1 = y_1 - \alpha.^*$$

It follows from (8) and (9), putting $\mathrm{Z}=1$, that the contribution to a is

$$\frac{80e^6}{\pi^2\sqrt{3}hc\,(k\mathrm{T})}\cdot e^{\alpha-y_1}\,\vartheta \tag{40}$$

$$= \frac{80}{\pi^2\sqrt{3}}\cdot\frac{e^6\nu_1}{c}\,\frac{1}{y_1}\cdot\frac{e^{-y_1}}{(k\mathrm{T})^3}\cdot e^{\alpha}\,\vartheta. \tag{41}$$

We must now consider the contribution due to "bound-free" transitions from states of higher energy than those given by $l=1$ and $l=2$. Using Milne's result which was justified by Gaunt† as a good approximation for large total quantum number l, and not too high frequency of radiation, the rate of absorption of radiation of frequency ν by n_l l-electrons is given by

$$(a_\nu)_l = \frac{16\pi^2\mathrm{Z}^2e^6}{3\sqrt{3}ch^3}\,\frac{\nu_1}{\nu^3}\,\frac{n_l}{l^2}\,\Delta(l)\,\frac{e^{\alpha+h(\nu-\nu_l)/k\mathrm{T}}}{e^{\alpha+h(\nu-\nu_l)/k\mathrm{T}}+1} \tag{42}$$

where

$$\Delta(l) = \frac{1}{(l-\frac{1}{2})^2} - \frac{1}{(l+\frac{1}{2})^2}.$$

---

[*] *Cf.* Milne, 'Mon. Not. R. Astr. Soc.,' vol. 90, p. 769 (1930).

[†] 'Phil. Trans.,' A, vol. 229, p. 163 (1930).

Using this result for $l \geqslant 3$ we have for the contribution to the atomic absorption coefficient

$$a_l = \frac{\int_{\nu_l}^{\infty} (a_\nu)_l \, \nu^3 e^{-h\nu/kT} \, d\nu}{\int_0^{\infty} \frac{\nu^3}{e^{h\nu/kT}+1} \, d\nu}$$

$$= \frac{80}{\pi^2\sqrt{3}} \frac{Z^2 e^6 \nu_1}{c\,(kT)^3} \cdot \frac{n_l \Delta(l)}{l^2} \cdot e^{-\nu_l+\alpha} \int_{\nu_l}^{\infty} \frac{d\nu}{e^{\alpha + h(\nu-\nu_l)/kT}+1} \tag{43}$$

$$= \frac{80}{\pi^2\sqrt{3}} \cdot \frac{Z^2 e^6 \nu_1}{c\,(kT)^3} \frac{n_l \Delta(l)}{l^2} \cdot e^{-\nu_l+\alpha} \log(1+e^{-\alpha}) \tag{44}$$

This is again the classical result multiplied by a factor $e^{\alpha} \log(1+e^{-\alpha})$.

Using

$$\frac{n_l}{n_1} = l^2 e^{\nu_l - \nu_1},$$

and including the contribution from the electrons in quantum states $l = 1$ and $l = 2$ as calculated in §§ 3 and 4, it follows at once by comparison with C II, p 480 *et seq.*, that the total contribution from " bound-free " transitions to the absorption coefficient per normal hydrogen atom is

$$2\cdot 286 \frac{80}{\pi^2\sqrt{3}} \frac{e^6 \nu_1}{c} \frac{e^{-y_1}}{(kT)^3} \cdot e^{\alpha} \log(1+e^{-\alpha}). \tag{45}$$

We have therefore for the absorption coefficient per normal hydrogen atom

$$a = \frac{80}{\pi^2\sqrt{3}} \cdot \frac{e^6 \nu_1}{c} \cdot e^{\alpha} \left[\frac{\vartheta}{y_1} + 2\cdot 286 \log(1+e^{-\alpha}) \right. \quad (kT)^3 \tag{46}$$

and for the mass absorption coefficient

$$\kappa = \frac{a n_1}{\rho} = \frac{80}{\pi^2\sqrt{3}} \frac{e^6 h^2}{c\,(2\pi m)} [\vartheta + 2\cdot 286 y_1 \log(1+e^{-\alpha})] \frac{N_e N_i}{S(kT)^{\frac{1}{2}} \rho}, \tag{47}$$

where the first term in square brackets represents the contribution from " free-free " and the second that from " bound-free " transitions A comparison of columns (3) and (5) in Table I shows that the ratio of " bound-free " to " free-free " absorption increases slightly with increasing degeneracy in the range with which we are concerned, $(\infty > \alpha > -1)$ For values of y_1 greater than 3 the " bound-free " absorption is predominant In hydrogen the condition

$$y_1 > 3$$

involves

$$T < 0\cdot 5 \,.\, 10^5.$$

§ 6. *Extrapolation for the Absorption Coefficient in general.*—A similar process to that employed in C II (§ 9) leads to an expression

$$a = \frac{40}{\pi^4\sqrt{3}}\,\frac{e^2h^4}{cm\,(2\pi m)^{\frac{3}{2}}}\left\{\Sigma\,\chi_r\left[\vartheta + \frac{2\cdot 286\,\chi_r}{k\mathrm{T}}\log\left(1+e^{-\alpha}\right)\right]x_{r+1}\right\}\frac{\mathrm{N}_e}{\mathrm{S}(k\mathrm{T})^{\frac{3}{2}}}, \quad (48)$$

for the atomic absorption coefficient in general, where

x_r = fraction of total number of atoms which have lost $(r-1)$ electrons.
χ_r = rth ionization energy

This formula is applicable when $\chi_r > 3k\mathrm{T}$ and $\alpha > -1$, *i.e.*, for incipient degeneracy. For extreme degeneracy, as we have seen, the ionization formula cannot be applied except for very low temperatures. For high degeneracy when the ionization formula cannot be applied the contribution from the " bound-free " transitions is small, and moreover thermal conduction becomes of more importance than the radiative transfer of energy.*

In the range of values considered for α and χ_r it is clear from the values of ϑ, $\log(1+e^{-\alpha})$ and S in Table I that no very great error is incurred by using the classical formula. It appears therefore that the " standard model " of a star gives such values for the density and temperature at the interface, as to justify the use of the classical formula up to the interface and to make the use of the " degenerate " formula inside the interface imperative. The results obtained here apply to the " straight mean " absorption coefficient. From fig. 2, which gives the behaviour of the Rosseland mean of the absorption due to " free-free " transitions, we see that with increasing degeneracy the Rosseland mean falls off much more slowly from the value which would be computed from the classical formula. It would therefore, probably be equally justifiable to use the classical expression for the Rosseland mean in the range considered.

It is a pleasure to express my thanks to Professor E. A. Milne for his helpful criticism during the course of this work.

Summary.

Expressions have been found by various writers for the absorption coefficient of a classical gas and a highly degenerate gas. In the present paper the absorption coefficient of a gas in intermediate states is investigated. Values are found for the " straight mean " absorption coefficient due to both " bound-free " and " free-free " transitions, and for the opacity, considering " free-free " transitions only. The application of the results to the two-phase " standard model " of a star is discussed.

---

* See Kothari, 'Mon. Not. R. Astr. Soc.,' vol. 93, p. 62, *et seq.* (1932).

Adsorption, Oriented Overgrowth and Mixed Crystal Formation.

By C. W. Bunn.

(Communicated by C. N. Hinshelwood, F.R.S.—Received March 21, 1933.)

[Plate 8.]

Introduction.

In this paper is considered the cause of strong adsorption of certain substances on specific crystal faces, and the relations between this phenomenon and the other two types of phenomena set down in the title—oriented overgrowth of different crystals on each other, and mixed crystal formation between two substances. A theory of adsorption is suggested, according to which this mode of association of two substances is related to the other two modes, oriented overgrowth and mixed crystal formation, very much more closely than is commonly realized.

Present knowledge of strong adsorption on specific crystal faces is due to the study of habit modification by dissolved impurities. Urea nitrate grown from a solution containing methylene blue forms a good example; Gaubert* found that the presence of methylene blue in the solution modifies the habit of the crystals, and at the same time it becomes included in the crystals, colouring them blue. The modification of habit in such cases appears to be caused by strong preferential adsorption of the impurity on certain crystal faces; the rate of growth of these faces (thickness of solid deposited in a given time) is reduced, and they are therefore prominent on the final crystal, since it is always the most slowly growing faces which predominate. The evidence for this view is that when the adsorbed substance is a coloured substance and can thus be seen in the crystal, as in the above example, it is seen to be located in pyramid-shaped regions opening out from the centre of the crystal to the faces whose rate of growth has been reduced. Other typical examples with coloured impurities are: urea nitrate in presence of picric acid,* which affects different faces from those affected by methylene blue, $Pb(NO_3)_2$ in presence of methylene blue†, and alum in the presence of diamine sky blue FF.‡

---

* 'C. R. Acad. Sci. Paris,' vol. 143, p. 936 (1906).

† Buckley, 'Z. Kristallog.,' vol. 76, p. 147 (1930).

‡ Keenan and France, 'J. Amer. Ceram. Soc.,' vol. 10, p. 821 (1927).

The same appears to be true for colourless impurities. Three examples are described below in Section I—sodium chlorate in presence of sodium thiosulphate,[*] potassium sulphate in presence of potassium chlorate,[†] and potassium sulphate in presence of potassium nitrate.[†]

It appears likely, judging from these examples and from the work of Gaubert[‡] on the effect of a large number of substances on ammonium chloride crystals, that whenever crystal habit is modified by the presence of an impurity in the solution, the impurity becomes included in the crystal owing to strong preferential adsorption on the faces whose rate of growth is reduced.

Gaubert[§] in 1925 suggested that habit modification is due to the same causes as oriented overgrowth , the discussion leading up to this suggestion does not seem convincing, but he gave very good evidence in support of it, in the fact that $Pb(NO_3)_2$ forms oriented overgrowths on methylene blue crystals.

Sloat and Menzies,[||] acting on Gaubert's suggestion, attempted to obtain oriented overgrowths of urea on octahedral faces of sodium chloride, failed, and abandoned the idea.

Buckley (in various papers) has attempted to explain adsorption on specific faces, for potassium salts of oxy-acids, as due to adherence of an ion of impurity by a pair or a triangle of its oxygen atoms to a pair or triangle of oxygen atoms of an ion exposed on the growing face of the crystal. This idea seemed to work sometimes, but failed at others.

I have approached the problem by examining the properties of modified crystals containing impurity. It is probable that the arrangement of the particles of impurity in these crystals is the same as it was on the faces of the crystals during growth; consequently, if the nature of these impure crystals were understood, we should know something about the arrangement of the adsorbed particles on the growing crystal faces. Gaubert (*loc. cit.*, 1918, 1925) has already stated, without giving much evidence, that these impure crystals (which will be called here " adsorption bodies ") are solid solutions in which both components are oriented, *i.e.*, mixed crystals. In the following pages, the properties of " adsorption bodies " and mixed crystals are compared; a good deal of new evidence is presented.

---

[*] Buckley, 'Z. Kristallog.,' vol. 75, p 15 (1930).

[†] Buckley, 'Z. Kristallog.,' vol 81, p 157 (1932).

[‡] 'Bull. Soc. Franc. Min.,' vol. 38, p 149 (1915).

[§] 'C R. Acad Sci Paris.,' vol. 167, p. 491 (1918); vol. 180, p 378 (1925)

[||] 'J. Phys. Chem.,' vol. 35, p. 2005 (1931).

I.—*Mixed Crystals and "Adsorption Bodies."*

Fine Structure.—X-ray methods have not, up to the present, given any direct evidence of the structure of adsorption bodies Powder spectra of two examples were obtained by France and collaborators,[*] but in each case the spectrum was the same at that of the pure substance , the amount of impurity present in the crystals was too small to cause any appreciable modification of the spectra

Optical properties show, firstly, that a contaminated region of an adsorption body (in the examples studied below) is, as far as microscopic examination is able to decide, a single solid phase when first formed The units of impurity are too small to be seen individually, even under high powers.

Secondly, optical properties give evidence of orientation of the units of impurity. Interpretation of the evidence is simplest for isometric crystals. Cubic crystals of $Pb(NO_3)_2$ or $Ba(NO_3)_2$ coloured by methylene blue are birefringent and pleochroic. Each cube is divided diagonally into six pyramids with their apices at the centre of the crystal and opening out to the six cube faces, each pyramid having the same optical orientation with respect to its own cube face The whole structure is a sort of mimetic twinning structure, fig 1. Now pleochroism might be produced in a coloured crystal by strain , but pleochroism due to strain is not likely to assume the precise regular character observed, and therefore this possibility can be ignored. The type of regularity observed is evidence of orientation of the impurity with respect to the crystal lattice , further, the impurity is oriented in a particular way with respect to the cube face on which it is deposited. Gaubert (*loc. cit.*, 1918) concluded that pleochroism is evidence of orientation of the impurity, though he does not mention the " twinning " structure.

Pleochroism has also been observed by Tammann and Laass[†] and Gaubert[‡] in certain organic crystals coloured by dyes, grown either from alcoholic solution or from melts.

Orientation of colourless impurities in adsorption bodies is sometimes indicated by changes in the birefringence of the crystals. Interpretation is simplest for isometric crystals containing birefringent impurities. Buckley (*loc. cit*, 1930) found that sodium chlorate, which normally crystallizes in cubes, forms tetrahedra in a solution containing thiosulphate, and that the modified crystals contain some thiosulphate. I have observed that when a strongly supersaturated solution of sodium chlorate containing thiosulphate crystallizes rapidly on a microscope slide, the crystals are weakly birefringent, and that crystals having both cube and tetrahedron faces show the strongest birefringence in sectors opening out from the centre of the crystal to the tetrahedron faces The thiosulphate is evidently located chiefly in these regions A pure tetrahedron (*i.e.*, with no cube faces) consists of four pyramids with their apices at the centre of the crystal, fig. 2, each pyramid having the same optical orientation with respect to its own tetrahedron face Again we have a sort of mimetic twinning structure, analogous to that of barium nitrate cubes containing methylene blue, except that there are four sectors instead of six; again we have evidence of orientation of the thiosulphate units with respect to the face on which they are deposited.

---

[*] Foote, Blake and France, 'J. Phys. Chem.,' vol. 34, p. 2236 (1930); Weinland and France, 'J. Phys. Chem ,' vol 36, p. 2832 (1932).

[†] 'Z. anorg. Chem ,' vol. 172, p. 65 (1928).

[‡] 'C. R. Acad. Sci. Paris,' vol. 194, p. 109 (1932).

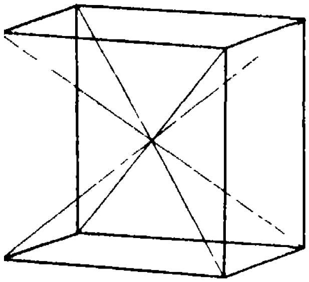


FIG. 1.—" Twinning " structure of $Pb(NO_3)_2$ cube containing methylene blue, and of mixed crystals of NH_4Cl with metallic chloride dihydrates.

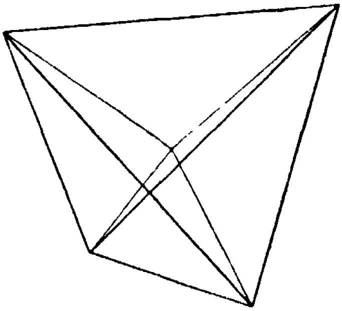


FIG 2 " Twinning " structure of $NaClO_3$ tetrahedron containing thiosulphate.

The optical structure of these crystals is not always so simple as that described. Crystals which start as cubes show a diagonal division into very weakly birefringent sectors, exactly like cubes of barium nitrate containing methylene blue, evidently thiosulphate is adsorbed and retained on cube faces, though far less strongly than on tetrahedron faces. When tetrahedron faces subsequently develop on these crystals, the optical structure is naturally rather complex.

A good example involving a crystal belonging to one of the less symmetrical systems is that of potassium sulphate (orthorhombic) crystallized in presence of potassium chlorate. Buckley (*loc. cit.*, 1932) found that this impurity modifies the habit of K_2SO_4 crystals, the chief effect being on 001 faces. I have observed that in crystals of K_2SO_4 grown rapidly on a microscope slide in presence of $KClO_3$, the sectors ending in the 001 faces are more weakly birefringent than the rest of the crystal, when observed from a direction normal to 010, fig. 17, Plate 8; even in ordinary transmitted light, diagonal dividing lines are visible, as in fig. 16, Plate 8 The $KClO_3$ is evidently adsorbed and retained by the 001 faces of K_2SO_4 in such a way as to reduce the birefringence seen from this direction.

The same effect is observed when the impurity present in solution is KNO_3; but the appearances are often more complex, owing to twinning of the K_2SO_4 crystals.

In these adsorption bodies, therefore, the impurity units are oriented with respect to the crystal lattice, and they are very small, beyond microscopic size. A fine structure of this type is, as far as optical properties are able to decide, that of a mixed crystal.

Other properties of mixed crystals and adsorption bodies will now be considered. The evidence furnished by these properties is not direct evidence of structure, but the analogies presented give strong support to the view that the two types of crystals are essentially similar.

"*Twinning*" *Structure.*—We have seen that in adsorption bodies bounded by faces all of the same form, oriented deposition of the impurity on the faces gives rise to a sort of mimetic twinning structure. Mixed crystals of ammonium chloride with metallic chloride dihydrates* have structures entirely analogous to that of barium nitrate containing methylene blue. If solutions of ammonium chloride containing the divalent chlorides of manganese, iron, cobalt, nickel or copper are crystallized, cubes (often with curved faces) are at first formed; these cubes are birefringent, and are divided diagonally into six sectors opening out from the centre of the crystal to the cube faces; quartz wedge and conoscopic observations show that the optical orientation of a sector with respect to its own cube face is the same for all sectors—each sector is a uniaxial crystal with its optic axis normal to the cube face.

These mixed crystals are usually described as "anomalous" because the substances forming them are not isomorphous. The similarity of the "twinning" structures to that of $Ba(NO_3)_2$ crystals containing methylene blue suggests that adsorption bodies are of the same nature as mixed crystals of non-isomorphous substances.

Factors determining Composition.—The amount of impurity in an adsorption body depends, not only on the nature of the face on which deposition occurs, and on the concentration of the impurity in the solution or melt,* but also on the rate of growth of the crystals. In the examples studied, the optical effects showing the presence of impurity are stronger when the crystals grow rapidly than when they grow slowly. Tammann and Laass (*loc. cit.*) observed the same thing in benzophenone crystals grown from melts containing dyes. Thus crystals when growing slowly are able to reject the impurities arriving at their faces, whereas when growing rapidly they are unable to do this so completely.

---

* Clendinnen and Rivett, 'J. Chem. Soc.,' vol. 119, p. 1329 (1921), vol. 121, p. 801 (1922); vol. 123, pp. 1344, 1634 (1923).

Mixed crystals of the isomorphous substances $(NH_4)_2SO_4$ and K_2SO_4 behave in a similar way. In the system $(NH_4)_2SO_4$-K_2SO_4-H_2O, at equilibrium the ratio $(NH_4)_2SO_4 . K_2SO_4$ in the solid phase is always far less than that in the solution. Crystals grown slowly reject the excess $(NH_4)_2SO_4$ which arrives at their faces, and have a composition approaching the equilibrium ratio Crystals grown rapidly in a thin film of solution on a microscope slide have the appearance of figs. 18 and 19, Plate 8 ; the rapidly growing end faces are unable to reject all the excess $(NH_4)_2SO_4$, and therefore the sectors ending in these faces contain more $(NH_4)_2SO_4$ and are more strongly birefringent (since the birefringence of $(NH_4)_2SO_4$ is much greater in all the directions than that of K_2SO_4) than the sectors ending in the more slowly growing side faces, which are able to reject more of the unwanted $(NH_4)_2SO_4$. That the end sectors are more unstable with respect to the solution than are the side sectors is shown by the fact that the end sectors redissolve in the mother liquor in a few minutes, these regions becoming quite hollow. Similar effects were observed in mixed crystals of K_2SO_4 and K_2CrO_4, and $(NH_4)_2SO_4$ and K_2CrO_4. The essential similarly of these crystals to that of K_2SO_4 grown in presence of $KClO_3$ is well illustrated by the similarity of figs. 19 and 17, Plate 8. The difference lies in the fact that in the latter system the equilibrium concentration of $KClO_3$ is solid K_2SO_4 is near zero.

In general the chief difference between adsorption bodies and ordinary mixed crystals is that in the former the amount of impurity taken up depends in a highly specific way on the nature of the face ; in the latter it does not, unless the faces grow at different rates. Even this difference, however, is only a difference of degree ; for the monoclinic mixed crystals of hydrated metallic sulphates $MSO_4 . 7H_2O$ (where M = Fe, Mn, Co, Cu, etc), which when grown rapidly on microscope slides are flat rhombs with inclined extinction, are divided diagonally into sectors, one opposite pair having a slightly different extinction angle and therefore a different composition from the other pair. Since the faces concerned grow at about equal rates, it appears that the ease of rejection of unwanted material depends to a certain extent on the nature of the face as well as on its rate of growth. Fig. 20, Plate 8, shows a crystal containing manganous and copper sulphates; the irregular portion growing out rapidly on the left has an extinction position different again from those of the pairs of sectors. The strongest effects of this type were observed with the combinations Fe-Cu (4 : 1), Mn-Cu (1 1), and Ni-Cu (1 1 and 3 : 1), weak effects with Zn-Fe (1 : 1 and 3 : 1); and no effect with Mg-Zn (1 : 1 and 3 . 1), Fe-Mn (4 : 1), Cd-Cu (1 : 1)

---

* Gaubert (*loc. cit.*, 1906) ; Tammann and Laass (*loc. cit*), Gaubert, ' C. R. Acad. Sci Paris,' vol. 147, p. 632 (1908).

Stability.—Rapidly grown adsorption bodies are unstable in their mother liquor. Sodium chlorate crystals containing thiosulphate partially redissolve and become practically isotropic within one day Potassium sulphate crystals containing chlorate change more slowly, but within one day the 001 faces become etched and hollowed. The analogous dissolution of the rapidly growing sectors of $(NH_4)_2SO_4$–K_2SO_4 mixed crystals has already been mentioned

More important than this is the fact that some adsorption bodies in the dry state separate into two solid phases Ammonium chloride cubes grown in a solution containing much urea (but not saturated with urea) take up very little urea, for the crystals are only very weakly birefringent, but within a few minutes of their formation, thin birefringent laminæ parallel to the cube faces appear, a few to each crystal; presumably these consist of urea which has separated from the composite crystal. Migration of urea through the crystal must occur in order to form these laminæ. It is evident that the amount of urea which can be held in solid solution in ammonium chloride is very small

Sodium chlorate crystals containing thiosulphate become practically isotropic in two days, and minute, strongly birefringent crystals, presumably of thiosulphate, appear on the surfaces Potassium sulphate crystals containing chlorate show, within a week, very small birefringent crystals in the sectors opening out to the 001 faces, and the general birefringence of these sectors increases.

Similarly, Tammann and Laass (*loc. cit.*) observed that the dye in their pleochroic benzophenone crystals grown from melts subsequently separated in the form of inclusions, leaving the body of the crystal colourless.

Certain mixed crystals behave in exactly the same way. Grimm and Wagner* found that mixed crystals of $BaSO_4$ and $KMnO_4$ containing up to 40% $KMnO_4$ are stable; mixed crystals containing up to 60% $KMnO_4$ can be obtained, but they rapidly become opaque owing to separation of minute $KMnO_4$ crystals.

---

* 'Z. phys. Chem.,' vol. 132, p. 131 (1928).

Mixed crystals of NH_4Cl with metallic chloride dihydrates also separate into two solid phases in the dry state. Specimens containing each of the five metals Mn, Fe, Co, Ni and Cu were prepared about three years ago and left either in the dry state or mounted in Canada balsam on microscope slides. In all specimens small birefringent crystals appeared on the surfaces, and the birefringence of the main crystal decreased practically to zero; this happened most rapidly (within a few months) with crystals containing Cu, and the small birefringent crystals were identified as crystals of $2NH_4Cl . CuCl_2 . 2H_2O$. Crystals containing the other metals changed more slowly, but are now practically isotropic.

The similarity between these "supersaturated" mixed crystals and adsorption bodies suggests that the latter are also "supersaturated" mixed crystals. This term, however, implies a definite, if small, equilibrium solubility In the only two systems in which the equilibrium relations have been worked out, K_2SO_4-KNO_3-H_2O* and NaCl-urea-H_2O,† no stable ranges were encountered, though it must be remembered that any stable ranges would be so small that they would probably not be detected unless they were being definitely sought. If the equilibrium solubilities are zero (if there is such a thing as complete immiscibility of solids), adsorption bodies should be described as "unstable mixed crystals."

II.—*Theory of Adsorption.*

Application of Knowledge on Mixed Crystal Formation to Adsorption.—Since adsorption bodies and mixed crystals appear to differ only in degree, existing knowedge on mixed crystal formation may be applied, in a suitably modified form, to the adsorption problems presented by the phenomena of habit-modification, and possibly to adsorption on crystal faces in general.

The most well-known examples of mixed crystals are those formed between isomorphous substances which have atomic lattice structures similar both in form and in absolute dimensions. According to Goldschmidt,‡ mixed crystal formation only takes place when the corresponding interatomic distances in the two separate crystals do not differ by more than about 10%; when this condition is fulfilled, the two substances form an extensive, or sometimes a complete, range of stable mixed crystals.

---

* Hamid, 'J. Chem Soc.,' p. 199 (1926).

† Ritzel, 'Z. Kristallog.,' vol. 49, p. 152 (1911).

‡ 'Trans. Faraday Soc.,' vol. 25, p. 268 (1929).

The reason why there is no appreciable stable mixed crystal formation between the substances which form adsorption bodies is that there is no overall similarity of lattice structures; but the fact that unstable mixed crystals are formed suggests that the lattices have a partial, very limited, similarity; we wish to know the nature of this partial similarity.

The Condition for Strong Adsorption.—Since the condition for stable mixed crystal formation is complete similarity of lattice structure in all respects, it seems reasonable to suppose that the condition for strong oriented surface adherence and formation of unstable mixed crystals on certain faces only is similarity of the lattice structures of the two substances along certain specific planes only, the rest of the lattice structures being dissimilar. In other words for stable mixed crystal formation a three-dimensional similarity of lattice structure is required, while for strong oriented surface adherence a two-dimensional similarity is required. Thus, if in two crystals the atomic arrangement on a particular plane of one crystal is similar to that on a particular plane of another crystal, and the interatomic distances are similar (within, say, 10% if the analogy with mixed crystals holds good), then one of the substances will adhere strongly to the appropriate crystal face of the other, and *vice versa.*

The conception of "similarity" needs closer examination. The structure of crystals in general can be regarded, to a first approximation, as a consequence of the necessity of packing spherical atoms of specific sizes, or groups of such atoms, in an orderly manner. The structures of mixed crystals can be regarded in the same way, though there is apparently a certain amount of distortion of the atoms to enable them to accommodate themselves to an intermediate lattice structure. There is no reason to suppose that conditions are any different for adsorbed substances on crystal faces, for we have seen that adsorption bodies and mixed crystals differ only in degree. We are therefore led to suppose that the similarity of lattice structures along specific planes must be such that a packing effect can be achieved, projecting atoms on the crystal face fitting into hollows of the thin layer crystal of adsorbed substance. This, of course, is only a crude picture, but it may be useful in the same way as the conception of the packing of spheres is useful in dealing with crystal structures in general.

This view seems inherently reasonable. If we imagine adsorbed particles more or less at rest on crystal surface, we should expect the most stable arrangement to be that in which the particles are arranged in the same way as they would be in a crystal of the substance; and we should expect the strongest adherence in crystals where this arrangement corresponds to the arrangement of particles on the crystal face in such a way that a packing effect can be achieved. The packing effect permits a closer average approach between atomic centres, and therefore a stronger binding, than would otherwise be possible There are, of course, other factors which will probably have to be taken into account ultimately, but for the present attention will be confined to the conception in its simplest form.

The adsorbed layer need not be continuous. Sometimes, as in the growth of modified crystals in solutions containing very small amounts of impurity, there may be isolated molecules or ions of impurity here and there on the growing crystal face. Perhaps the most satisfactory conception of the necessary condition for strong adsorption, covering these systems as well as continuous layers, is this . When the crystal face has a structure similar to that of a face of a crystal of the adsorbed substance, a particle of the latter will adhere to the crystal face just as it would to the corresponding face of its own crystal, since it experiences similar forces

The condition postulated as necessary for strong adsorption of a substance on a crystal face of another appears to be identical with that known to be necessary for oriented overgrowth of a crystal of one substance on a crystal face of another. Of course, it may be that the necessary degree of closeness of interatomic distances for adsorption is different from that required for oriented overgrowth; but it should at any rate be possible, in cases where very strong adsorption is known to occur, to obtain oriented overgrowths; that is to say, when it is known that a certain substance in solution has a strong effect in modifying the habit of a crystal, it should be possible under suitable experimental conditions to obtain oriented overgrowths of the substance on those faces of the crystal whose rate of growth it depresses

It is hardly surprising that consideration of the properties of adsorption bodies had led to the idea that the essential condition for strong adsorption is qualitatively the same as that necessary for oriented overgrowth; for it seems obvious that adsorption is a necessary preliminary to the formation of an oriented overgrowth, and in fact an oriented overgrowth is simply an adsorbed layer which has been " developed " up to visible size.

Habit Modification --The reduction of the rate of growth of specific faces when impurities are strongly adsorbed on them can be explained without making any assumptions as to the mechanism of deposition of particles on the faces.

A certain amount of impurity is always included in those regions of the crystal built up by deposition on the affected faces, forming an unstable mixed crystal. The fact that the mixed crystal is unstable gives the explanation. On the growing face, particles of impurity are strongly adsorbed, together with particles of the main substance; when more material is deposited on these, an unstable mixed crystal is formed, which, being unstable, tends to redissolve, the net result being that the rate of growth is lower than it would be in absence of the impurity. Looking at it in another way, there is an attempt to build up a pure crystal, since the mixed crystal is unstable; but the impurity is adsorbed so strongly that it is difficult to eject it, and it is the conflict between these forces that is responsible for the reduction of the rate of growth.

There is some evidence in favour of the view that a redissolving process occurs concurrently with growth. When sodium chlorate crystallizes on a microscope slide in presence of a large concentration of thiosulphate, the first crystals grow as beautifully rounded tetrahedra , later, as growth slows down, the crystal faces become plane. Also, in presence of a comparatively small concentration of thiosulphate, when the crystals have both cube and tetrahedron faces, the tetrahedron faces are often curved at first, while the cube faces are perfectly plane. The rounded form suggests a dissolving process, which is most urgent when large amounts of impurity enter the crystal owing to rapid growth.

The greatest modification of habit would be expected where there is very strong similarity of specific lattice planes, but complete dissimilarity of the rest of the structures. As regards dissimilarity of the complete structures, this seems to be true; when complex organic substances such as dyes modify inorganic crystals, very small quantities are usually sufficient to cause considerable modification of habit

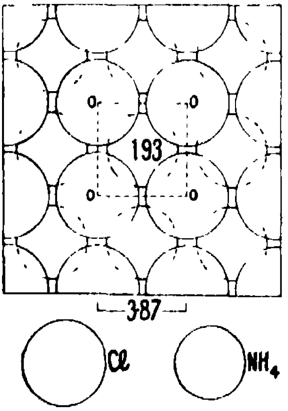
The weakest modification of habit would be expected where there is either little correspondence of lattice structures on any planes, or else if the whole structures are similar. In the latter case, strong adsorption occurs on all faces, and since the mixed crystal formed is stable there is no urgent necessity for rejection of the foreign particles, though a certain amount of readjustment of composition may be necessary to attain the equilibrium composition of the mixed crystal. This agrees with the facts; mixed crystals of isomorphous substances usually do not differ very much in habit from the pure components.

Before considering typical examples of habit modification in the light of these conceptions, one intermediate between typical mixed crystal formation and adsorption will be examined.

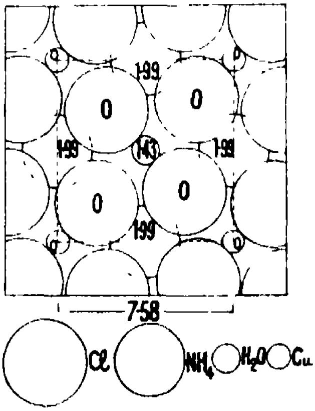
III.—NH_4Cl *and* $2NH_4Cl . CuCl_2 . 2H_2O$.

In the system $NH_4Cl–CuCl_2–H_2O$ (Clendinnen and Rivett, *loc. cit.*), NH_4Cl takes up $CuCl_2 . 2H_2O$ to form a small range of mixed crystals containing up to about 1·5% $CuCl_2$ at 25° C. There is a compound $2NH_4Cl . CuCl_2 . 2H_2O$ with a small range of mixed crystals on either side of it We may regard the mixed crystals containing up to 1·5% $CuCl_2$ as mixed crystals formed between NH_4Cl and $2NH_4Cl . CuCl_2 . 2H_2O$ It has already been mentioned that these mixed crystals are cubes and have the "twinning" structure characteristic of adsorption bodies bounded by only one type of face.

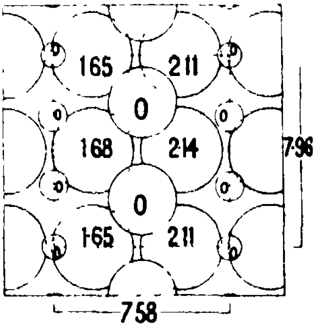
The atomic lattice structure of $CuCl_2$ $2H_2O$ is not known, but that of $2NH_4Cl . CuCl_2 . 2H_2O$ has been worked out.* It is tetragonal, with unit cell dimensions $a = 7·58$ A., $c = 7·96$ A, c being only slightly greater than a. NH_4Cl is body-centred cubic, with unit cell edge $a = 3·87$ A. Now twice 3·87 is 7·74, which is midway between the lengths of the a and c edges of $2NH_4Cl$ $CuCl_2 . 2H_2O$. If the structures of the two substances are compared, it can be seen that the structure on the basal plane 001 of $2NH_4Cl . CuCl_2 . 2H_2O$, fig. 4, is very similar to that on the cube face of NH_4Cl, fig 3; the structure on the 100 face of the double salt, fig. 5, is not very similar to that of the cube face of NH_4Cl, except in the correspondence of the unit cell



3



4



5

FIG. 3.—Atomic arrangement on cube faces of NH_4Cl.
FIG. 4.—Atomic arrangement on 001 faces of $2NH_4Cl$ $CuCl_2 . 2H_2O$.
FIG. 5.—Atomic arrangement on 100 faces of $2NH_4Cl$ $CuCl_2 . 2H_2O$.

The figures within the diagrams denote distances in A of atomic centres below the plane of reference.

---

* Wyckoff, "The Structure of Crystals," 1931.

edges, *c* of the double salt being almost exactly double *a* of NH_4Cl. Thus there is a close correspondence of the structure of the basal plane of the double salt with that of the cube face of NH_4Cl, the rest of the structures being only partly similar. It is no doubt the partial overall similarity of the two structures which is responsible for the formation of a small range of mixed crystals stable in solution. The close correspondence of the structure on the basal plane of the double salt with that on the cube face of NH_4Cl suggests an explanation of the " twinning " structure. Suppose the first nucleus is a cubic NH_4Cl crystal; on each cube face of this crystal, the double salt would deposit with its basal plane in contact with the cube face of NH_4Cl; one can imagine the NH_4 ions of the latter fitting into the hollows between Cl ions of the former, thus achieving the packing effect necessary for strong adsorption. That $2NH_4Cl . CuCl_2 . 2H_2O$ does orientate in this way on 100 faces of NH_4Cl is shown by the fact that oriented overgrowths are easily obtained, with 001 of the double salt in contact with 100 of NH_4Cl, oriented in exactly the expected manner. Further growth of the postulated NH_4Cl nuclei would cause the double salt to be included as mixed crystal, which is stable in contact with solution provided it does not contain more than 1·5% $CuCl_2$ The resulting crystal would have a " twinning " structure of exactly the type which in fact is formed, and each pyramid would be an optically negative uniaxial crystal with its optic axis normal to the cube face (since the double compound is optically negative); this is in fact the truth.

This case forms an admirable link between ordinary mixed crystal formation and adsorption. The formation of mixed crystals stable in solution is due to a partial overall similarity of lattice structure, while the formation of cube faces and the " twinning " structure can be explained in exactly the way which has been visualized for adsorption bodies, through strong similarity of the lattice structures on specific planes.

IV.—*Ammonium Chloride and Urea, Ammonium Bromide and Urea.*

NH_4Cl grows from pure solutions in the form of skeletal growths formed by very rapid growth along the axial directions. The presence of urea in the solution causes the formation of cubes of NH_4Cl. The effect is a very strong one, for one molecule of urea to every six of NH_4Cl is sufficient to cause the formation of perfect cubes even when crystallization takes place very rapidly, as in the rapid evaporation of a drop of solution on a microscope slide. When grown slowly in a day or two in a crystallizing bath, one molecule of urea to twelve of NH_4Cl is sufficient. The change from skeletal growth to formation of perfect cubes represents a very considerable reduction of rate of growth in the axial directions.

Structure.—Ammonium chloride is body-centred cubic with $a = 3{\cdot}87$ A. The arrangement of particles on a cube face is shown in fig. 3; all particles on a face are of one sort—either NH_4 or Cl ions—and their centres lie on a square network, the distance between neighbouring particles being 3·87 A.

Urea is tetragonal, and belongs to the tetragonal scalenohedral class. (Normal habit shown in fig. 8.) Its unit cell contains two molecules, and has dimensions $a = 5{\cdot}67$, $c = 4{\cdot}73$ A. (Wyckoff, *loc. cit.*). A projection on the basal plane, turned 45° from the usual position of representation for reasons which will appear directly, is shown in fig. 6. Now the points marked with small crosses form a square network of points 4·01 A. apart. (4·01 is the semi-diagonal of the unit cell area on the basal plane.) 4·01 and 3·87 differ by less than 4%, and the two structures would therefore fit together very well; moreover, a packing effect could be achieved by particles on the cube face of NH_4Cl fitting into the spaces marked with crosses, which are gaps between NH_2 groups of different urea molecules. Note that in order to fit the structures together, the urea structure must be placed so that its 110 faces are parallel to cube edges of NH_4Cl. Presumably 4% distortion

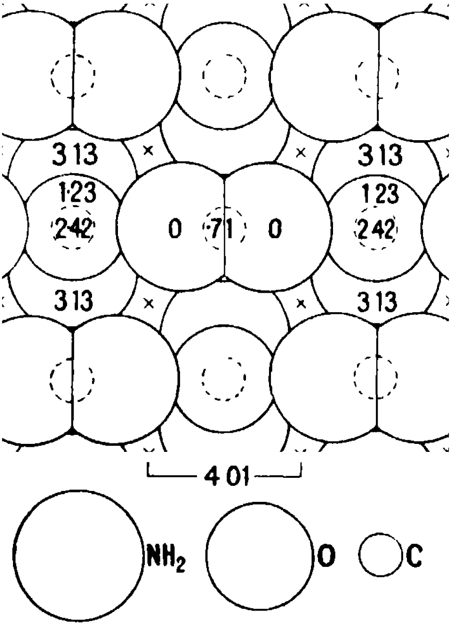


Fig. 6.—Atomic arrangement on 001 faces of urea (after Wyckoff).

is permissible, just as up to 10% distortion is permissible in mixed crystals.

Evidence that urea does orientate in this way on cube faces of NH_4Cl is provided by oriented overgrowths.

Oriented Overgrowth.—It has already been pointed out that if the present theory is true, it should be possible, where strong adsorption is known to occur, to obtain oriented overgrowths.

Oriented overgrowths of urea crystals on cube faces of NH_4Cl were easily obtained in the following way. Small cubes of NH_4Cl were grown on a microscope slide by allowing a drop of aqueous NH_4Cl solution containing 5% urea to evaporate in air The mother liquor was removed by filter paper, a drop of warm, strongly supersaturated solution of urea in equal parts of alcohol and benzene was added, and quickly covered by a cover-glass. Crystallization occurred immediately, and the NH_4Cl crystals were covered with urea crystals oriented in exactly the expected manner. Fig. 21, Plate 8, shows the freedom with which these oriented overgrowths are produced. Fig. 7 sums up the observations made. The urea crystals grow with their optic axes normal to NH_4Cl cube faces ; every NH_4Cl crystal shows, in convergent polarized light between crossed nicols, the uniaxial interference figure of urea, the optic axis being normal to the cube face. The basal plane of urea is thus in contact with the cube face of NH_4Cl. The 110 faces of urea are always parallel to cube edges. Note that on any particular cube face, the two possible orientations of urea crystals—with either 110 or $1\bar{1}0$ parallel to a particular cube edge—often occur together ; it appears, therefore, that oriented nuclei of urea are formed independently on different points of the same cube face.

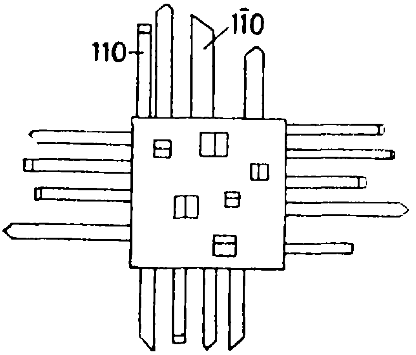
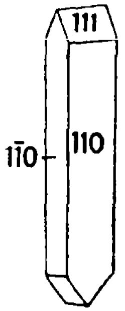


FIG. 7 —Oriented overgrowths of urea on cube faces of NH_4Cl.

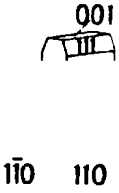
These oriented overgrowths can be obtained by using a purely alcoholic solution of urea; but they are not formed so readily, and the NH_4Cl crystals dissolve rather rapidly in the solution. Attempts to obtain oriented overgrowths from an aqueous solution supersaturated with urea and saturated with NH_4Cl failed.

Reciprocal Action.—It is an important consequence of the present theory that adsorption is a reciprocal phenomenon, *i e.*, if a substance A is adsorbed on a particular crystal face of B, then B will be adsorbed on a particular crystal face of A. Therefore habit modifications should be reciprocal; since urea reduces the rate of growth of cube faces of NH_4Cl, then NH_4Cl should reduce the rate of growth of 001 faces of urea.

When urea is grown in aqueous solution it forms exceedingly long acicular crystals without definite end faces. In alcoholic solution the habit is as shown



8.



9.

FIG. 8.—Normal habit of urea crystal.

FIG. 9.—Urea crystal grown in solution containing NH_4Cl.

in fig. 8. When NH_4Cl is present in the solution (nearly saturated with respect to NH_4Cl), basal planes appear, fig. 9, sometimes almost extinguishing the 111 faces.

The same thing can be observed on a microscope slide, *i e.*, under conditions of rapid growth. It is best observed with solutions in equal parts of alcohol and water. At first, when crystal growth is rapid, 111 faces are present on the urea crystals; subsequently, as growth slows down, basal planes appear and sometimes nearly obliterate the 111 faces.

The fulfilment of the predictions of oriented overgrowth and reciprocal action forms good evidence that the theory does represent the true state of affairs, at any rate in this particular example.

In conclusion it may be noted that the overall structures of NH_4Cl and urea are very dissimilar; it is only on one plane of each that a similarity exists.

This fits in with the idea already expressed, that the strongest habit modification effects would be expected in cases where there is maximum similarity on specific planes, but maximum dissimilarity over the whole lattices.

Ammonium Bromide and Urea.—Ammonium bromide has the same structure as ammonium chloride, and normally grows in the same skeletal form. It is well known that urea causes it to grow in cubes. The interatomic distance on the cube faces is 4·05 A., *i.e.*, very close to the 4·01 of urea.

It was found that NH_4Br has a much stronger effect on the habit of urea than NH_4Cl has. In alcoholic solution quite small quantities of NH_4Br cause the formation of basal planes which obliterate the 111 faces of urea crystals; even in aqueous solutions the same effect is easily observable on a microscope slide, and when the solution is nearly saturated with NH_4Br the urea crystals are very squat prisms terminated by basal planes.

Oriented overgrowths of urea on NH_4Br cubes were easily obtained by the same method as that used for NH_4Cl. The alcohol in a few seconds dissolves sufficient NH_4Br to cause the appearance of basal planes on the urea crystals

V.—*Sodium Chloride and Urea.*

Urea causes octahedral faces of NaCl to appear. The effect is less strong than that of urea on NH_4Cl, it requires one molecule of urea to every three molecules of NaCl to cause the formation of perfect octahedra without cube faces. It is to be expected, therefore, that the correspondence of the lattice structures will be less close than in the case of NH_4Cl.

There is one complication; in the aqueous system there is a compound $NaCl . CO(NH_2)_2 . H_2O$, the structure of which is not known. However, the following evidence shows that the production of octahedral faces of NaCl can be explained as being due to urea itself. The evidence of oriented overgrowths will be described first, before considering the structures.

Oriented Overgrowth.—Attempts to obtain oriented overgrowths on NaCl from alcoholic solutions of urea failed. This is in accordance with the experience of Sloat and Menzies (*loc. cit.*). These workers found, however, that in some crystals NH_4Cl could form oriented overgrowths from the vapour state on certain cubic crystals where none formed from solution. Urea vapour was therefore tried in the present experiment. NaCl crystals having both cube and octahedron faces were grown on a microscope slide from aqueous solution containing urea. The mother liquor was removed and the slide placed in a vessel in which urea was vapourized *in vacuo*. Oriented overgrowths were readily obtained, most freely on the octahedral faces but also on the cube faces The urea crystals were irregular in shape, and although some information on their orientation was obtained by optical properties, it was desirable to grow crystals with faces. This was done by immersing the slide in a slightly supersaturated solution of urea in alcohol; many of the urea crystals developed into well-formed crystals showing the usual faces (110 and 111).

It was found that on the octahedral faces the urea crystals were oriented with their 111 faces in contact with the NaCl face, the edge 111 : $\bar{1}11$ being parallel to the edge of the octahedron face. Six different orientations, of which only two are shown in fig. 10, may occur on the same face.

On the cube faces the urea crystals were usually oriented with their 110 faces in contact, and their optic axes along the diagonals of the cube face, as in fig 11 Four different orientations are possible. Also on a few rare occasions urea crystals with their basal planes in contact and 110 faces parallel to diagonals of the cube face were seen; examples of this are also shown in fig. 11.

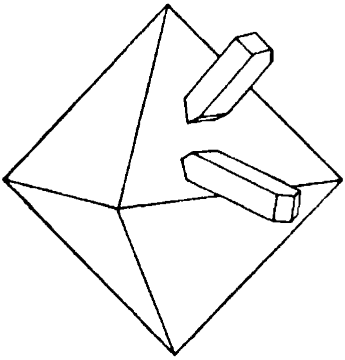


FIG 10.—Oriented overgrowths of urea on octahedron faces of NaCl On each face six different orientations may occur, of which only two are shown.

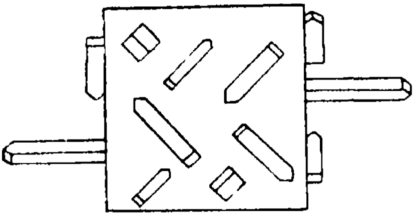
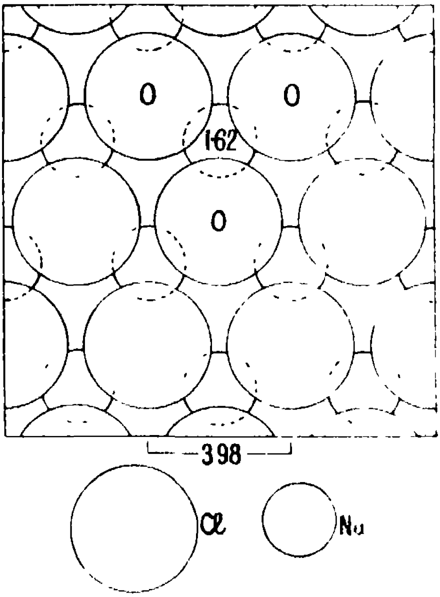


FIG. 11.—Oriented overgrowths of urea on cube faces of NaCl

Structure.—The reasons why these oriented overgrowths are formed can be seen from the structures on the faces concerned. The arrangement of particles on 111 of NaCl is shown in fig. 12; atomic centres lie at the corners of equilateral triangles with sides = 3·98 A The arrangement on 111 of urea is shown in fig. 13; only the molecules nearest the plane of reference are shown, the blank spaces being hollows into which NH_2 groups of the next set of urea molecules would fit. These hollows form isosceles triangles with one side 8·02 and the others 7·38 A—distances about double those of NaCl. But we can choose points such as those marked with crosses which lie at the corners of isosceles triangles with one side 4·01 and the others 3·69 A ; 4·01 differs from 3·98 by less than 1%, while 3·69 differs from 3·98 by 7 to 8%, and the structures therefore fit only moderately well The packing effect is also not very good, but it may be noted that if, say, Cl atoms are placed at the points marked, necessarily in a plane rather above the plane of reference, it is possible to put Na atoms in the hollows so that they occupy the correct positions with regard to the Cl atoms , possibly in the growth of 111 faces of NaCl this type of packing occurs



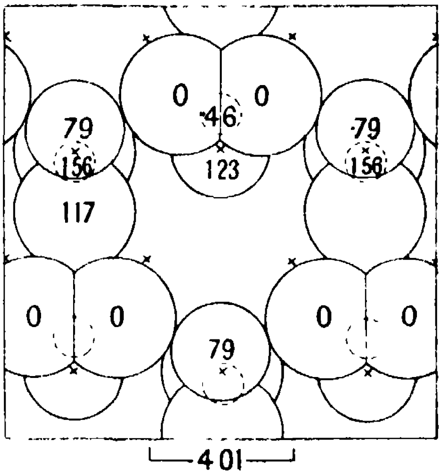
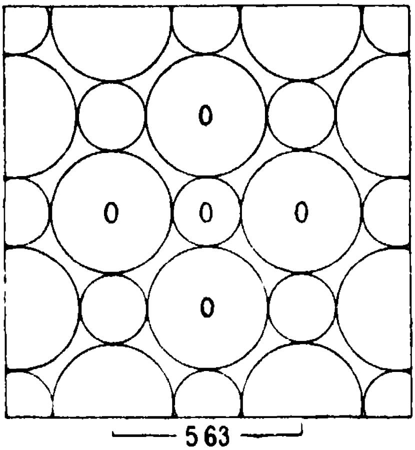


FIG 12 —Atomic arrangement on octahedron faces of NaCl. (on left)
FIG. 13 —Atomic Arrangement on 111 faces of urea (on right)

The arrangement on 100 of NaCl is shown in fig. 14 , the centres of Cl atoms form squares of side 3·98 A., with Na atoms in the centres of the squares On 110 of urea, fig. 15, the spaces between molecules form rectangles with one side 4·01 and the other 4·73 A. , this face would thus fit on 100 of NaCl with the optic axis of the urea crystal parallel to the diagonal of the cube face. 4·01 differs from 3·98 by less than 1%, while 4·73 differs from 3·98 by over 15% ; the structures therefore fit less well than do those of 111 of NaCl and 111 of urea ; also the structures do not pack at all well This is in accordance with the fact that oriented overgrowths on cube faces are formed less readily than on octahedron faces.

The rarely observed oriented overgrowths with 001 of urea in contact with

100 of NaCl and 110 faces of urea parallel to cube face diagonals are interesting and important. The arrangement on 001 of urea has already been given in fig. 6, 100 of NaCl in fig 14. There is a very close correspondence of unit cell dimensions, *a* for NaCl being = 5·63, and *a* for urea = 5·67, the difference being less than 1%. If oriented overgrowth were merely a question of correspondence of unit areas on specific faces, it would be expected that oriented overgrowths of this type would form more readily than the other types described ; and in addition it would be expected that adsorption of urea on cube faces of NaCl would be very strong, the rate of growth of these faces would be considerably reduced, and no habit modification of NaCl by urea would occur. If the structures are examined, it can be seen that although the unit areas correspond closely, a good packing effect cannot be achieved, and this may be the reason why oriented overgrowths of this type are rare, and adsorption on cube faces not so strong as on octahedral faces.



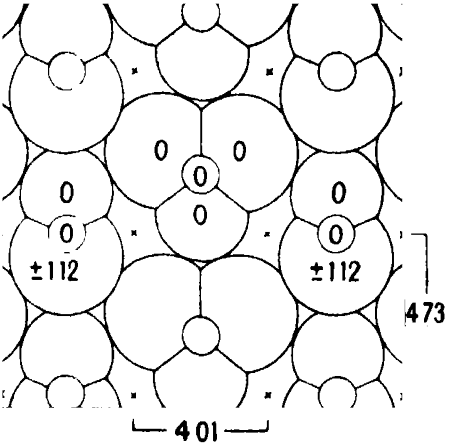


Fig. 14.—Atomic arrangement on cube faces of NaCl (on left).
Fig. 15.—Atomic arrangement on 110 faces of urea (on right).

The evidence provided by oriented overgrowths is thus unexpectedly varied ; but the fact that oriented overgrowths on octahedral faces are formed more readily than on cube faces fits in with the known effect of urea in reducing the rate of growth of the octahedral faces relatively to that of the cube faces.

No reciprocal modification of habit would be expected here, except a shortening of the acicular urea crystals ; for NaCl would be adsorbed fairly strongly on 111, less strongly on 110 (these two faces being the only forms on urea crystals), and only weakly on 001 of urea.

To investigate this, urea was crystallized from a solution in a mixture of 75% alcohol and 25% water, with NaCl present and without it. The urea crystals were normal in both cases, the length of the crystals was very variable, and no definite impression of shortening in presence of NaCl was gained; but the length of acicular crystals is usually so variable as to be of no great significance.

When urea was crystallized from a purely aqueous solution containing NaCl, the crystals were prisms bounded by basal planes only; there was no sign of any pyramid faces Since this only happens in aqueous solutions, and not in 75% alcohol solutions, it may be caused by adsorption of the compound $NaCl . CO(NH_2)_2 . H_2O$ on the basal plane of urea.

Retgers* found that urea has a much smaller effect on KCl than on NaCl, small octahedron faces being formed on the corners of KCl cubes. On KBr and KI urea has no effect at all. (All three potassium salts have the same space lattice as NaCl)

This is exactly what would be expected, for on the octahedron faces of KCl the interatomic distance is 4·44 A., *i.e.*, larger than on 111 of NaCl, and further removed from the spacing on 111 of urea; on octahedron faces of KBr and KI the spacing is greater still. This shows clearly the importance of atomic spacing, and the fact that the greatest effect occurs where there is closest correspondence of lattice structures is very good evidence in favour of the present theory.

VI.—*Other Examples.*

The following examples have been found in scientific literature. Where oriented overgrowths are known to be formed, this is in itself sufficient proof of the sort of similarity of lattice structure on specific planes which has been postulated as a necessary condition for strong adsorption.

$KClO_3$ *and* $KMnO_4$.—Buckley† found that $KMnO_4$ encourages 011 of $KClO_3$ with subsidiary effects on 110, $10\bar{1}$, 100 and 101. Also $KClO_3$ modifies $KMnO_4$, encouraging 102 and 001. He also observed that oriented overgrowths of $KMnO_4$ form readily on 011 of $KClO_3$, the 102 face of $KMnO_4$ being in contact with 011 of $KClO_3$.

This is a perfect example, as complete as that of NH_4Cl and urea, there is reciprocal action, and formation of oriented overgrowths with an affected face of each crystal in contact There appears to be a partial overall similarity of the lattice structures. It is no doubt this partial overall similarity of lattice structures that is responsible for the fact that there are habit modification effects on more than one face of each crystal ; and it is evidently 102 of $KMnO_4$ that is responsible for reducing the rate of growth of 011 of $KClO_3$, and *vice versa*

---

* 'Z. phys. Chem.,' vol. 9, p. 267 (1892).

† 'Z. Kristallog.,' vol. 82, p. 37 (1932).

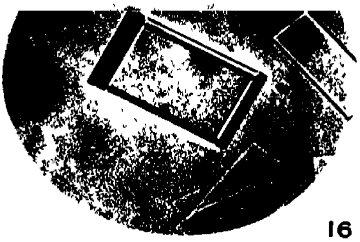
It seems at first sight rather surprising that oriented overgrowths are formed so readily, for the habit modification effects are not very strong (compare NaCl and urea). It may be that the partial overall similarity of lattice structures makes it easier for oriented overgrowths to form. But another view suggests itself. In view of the partial overall similarity of lattice structures, there may be a small stable range of mixed crystals. Now habit modification is usually less marked when stable mixed crystals are formed (see Section II) ; the comparative weakness of the modification of habit does not necessarily mean that adsorption is weak, it may mean that rejection is less urgent—the impurity can be retained in the crystal as a stable mixed crystal—and consequently there is less slowing down of the rate of growth and less habit modification than in crystals where the structures differ greatly. Thus adsorption may actually be strong, and therefore it is not so surprising that oriented overgrowths are formed readily

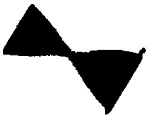
$Pb(NO_3)_2$ *and Methylene Blue* --Methylene blue encourages the cube faces of $Pb(NO_3)_2$. Gaubert (*loc cit* 1925) states that oriented overgrowths of $Pb(NO_3)_2$ form on methylene blue crystals , presumably it is the cube face of $Pb(NO_3)_2$ that is in contact with the methylene blue crystals, though this is not explicitly stated

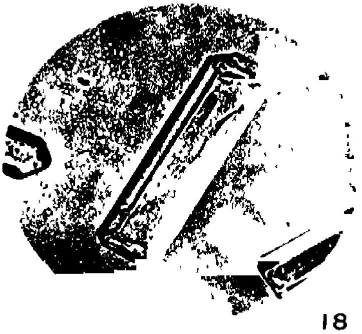
Urea Oxalate and Methyl Blue.—Gaubert* found that many dyes modify urea oxalate crystals ; some encourage 010, and one of these, methyl blue, forms oriented overgrowths readily on 010. He also states, as a general observation, that the coloured crystals of urea oxalate are solid solutions, but that "cristaux mixtes"—by which term he means (*loc cit* , 1918) the regular association of little crystals of visible size—can also be obtained, pre-presumably from solutions supersaturated with dye as well as with urea oxalate Gaubert's "cristaux mixtes" appear to be repeated oriented overgrowths, and if they are formed in all the examples mentioned, this would add many more to this list of crystals in which oriented overgrowths are formed on faces on which strong adsorption occurs. Unfortunately, Gaubert gives no details.

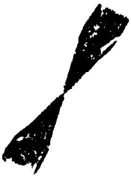
---

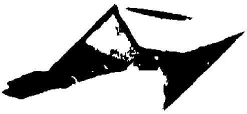
* 'C. R. Acad. Sci. Paris,' vol. 192, p 964 (1932)


16


17


18


19


20

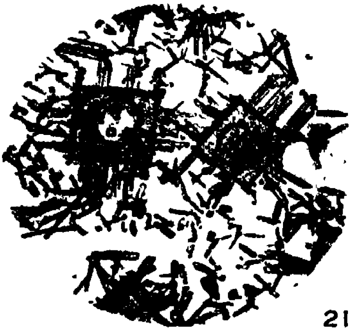

21

FIG. 16.—K_2SO_4 crystal grown in solution containing $KClO_3$. 330.

FIG. 17. The same crystal as (fig 16), between crossed nicols. 330.

FIG. 18. K_2SO_4–$(NH_4)_2SO_4$ mixed crystal, grown rapidly 240.

FIG. 19.—The same crystal as (fig 18), between crossed nicols. × 240.

FIG. 20—Mixed crystal from solution of manganous and copper sulphates, grown rapidly. 110.

FIG. 21 Oriented overgrowths of urea on NH_4Cl cube faces. (Some of the crystals are already detached, owing to dissolution of the NH_4Cl.) × 110.

(*Facing p* 588)

Together with the three examples given earlier, NH_4Cl, NH_4Br, and NaCl with urea, this makes at any rate six examples of oriented overgrowths where they would be expected according to the present view of adsorption, that is, six examples in which there is excellent evidence of similarity of specific lattice planes of crystals concerned in adsorption, and of mutual orientation of these planes on each other.

There are several other cases in which reciprocal habit modification occurs, but oriented overgrowths have not as yet been reported. These examples do not give direct evidence of similarity of lattice structures on specific planes, but they do suggest that adsorption is an affair concerning the crystal structures of the substances involved ; and it is probable that where only one face of each crystal is affected, it is these two faces which have similar structures.

We may expect some exceptions to the rule of reciprocal action, if the analogy between adsorption and mixed crystal formation holds good. For instance, sodium nitrate takes up a certain amount of sodium chlorate as mixed crystal, but sodium chlorate takes up no sodium nitrate The reason for this appears to be that $NaClO_3$ can exist as an unstable rhombohedral form, which often appears when a drop of $NaClO_3$ solution crystallizes on a microscope slide without cover glass ; it is this form which enters into $NaNO_3$ as mixed crystal, and is stabilized thereby But $NaNO_3$ has no unstable isometric form, and consequently the ordinary isometric form of $NaClO_3$ cannot take up $NaNO_3$ as mixed crystal. The same kind of thing may happen in habit modification, which may sometimes be due to an unstable form of the impurity which has faces with a lattice structure corresponding with that of particular faces of the modified crystals ; reciprocal action on the stable form of the impurity is not to be expected in such crystals.

Other exceptions are to be expected when on one of the substances adsorption occurs on faces which are already the principal or the only faces of the normal crystal. One such system –NaCl and urea—has already been encountered.

The following examples are taken from Buckley's work on potassium salts of oxy-acids.

K_2SO_4 *and* $KClO_3$.—$KClO_3$ encourages 001 and 010 (and 110 slightly) of K_2SO_4.* K_2SO_4 encourages 011 of $KClO_3$.†.

K_2SO_4 *and* KNO_3.—KNO_3 encourages 001 and also 010 of K_2SO_4.* I have observed that K_2SO_4 modifies the habit of KNO_3, but the precise nature of the modification has not been investigated. Buckley states† that he is studying modification of habit of KNO_3 and will no doubt include K_2SO_4 as one of the impurities.

---

* Buckley, 'Z. Kristallog.,' vol. 81, p. 157 (1932).

† *Ibid.*, vol. 82, p. 37 (1932).

$KMnO_4$ *and* $K_2Cr_2O_7$.—$K_2Cr_2O_7$ has a strong effect on 001 of $KMnO_4$. $KMnO_4$ also modifies $K_2Cr_2O_7$, but the precise nature of the modification is not known.*

K_2CrO_4 *and* $KClO_3$.—$KClO_3$ encourages 110 and 010 of K_2CrO_4 (*loc. cit.*). K_2CrO_4 encourages 011 of $KClO_3$.†

K_2CrO_4 *and* $KMnO_4$—K_2CrO_4 encourages 110 and 100 of $KMnO_4$ (*loc. cit.*). KNO_4 encourages 010, 102 and (weakly) 110 of K_2CrO_4 ‡

There is one doubtful case K_2SO_4 encourages 110 and 100 of $KMnO_4$ (*loc. cit.*). Buckley (*loc. cit.*) considers that $KMnO_4$ has no effect on the faces of K_2SO_4, but he states that it encourages the formation of twins and triplets on 110, and prevents the elongation along the *a* axis which often occurs in pure solutions, how much of this prevention of elongation is brought about by the twinning, which disturbs face development, and how much by adsorption effects, it is difficult to decide

There are thus five examples of reciprocal action, making with the three previous ones- NH_4Cl and NH_4Br with urea, and $KClO_3$ and $KMnO_4$-- eight in all; and two exceptions, one of which -NaCl and urea —is understood, and the other- K_2SO_4 and $KMnO_4$—doubtful

Finally, the effect of $K_2S_2O_6$ on K_2SO_4 may be mentioned Buckley (*loc. cit.) found that* $K_2S_2O_6$ *has a fairly strong effect on* 001 *of* K_2SO_4 Now K_2SO_4 is pseudo-hexagonal about the *c* axis, and the unit dimensions on 001 (Wyckoff, *loc cit*) are $a = 5·73$, $b = 10·00$ A On this face, half way up the unit cell as shown in the diagram in Wyckoff's book, there are K atoms occupying positions in a plane at the corners of isosceles (nearly equilateral) triangles, one side of which is 5·73 and the other two 5·76 A. $K_2S_2O_6$ is hexagonal, and it seems likely that it would be the basal plane which would fit on the to the pseudo-hexagonal basal plane of K_2SO_4 The crystal structure of $K_2S_2O_6$ has been determined by Helwig § The unit cell dimension $a = 9·82$ A., *i e.*, very close to *b* of K_2SO_4 The projection on the basal plane shows that there are K atoms occupying positions in a plane at the corners of equilateral triangles with sides 5·67 A. The structures on these planes would fit fairly well; the simplest way of regarding the adsorption of $K_2S_2O_6$ on

---

* 'Z. Kristallog.,' vol. 78, p. 412 (1931).

† *Ibid.*, vol 82, p. 37 (1932)

‡ *Ibid.*, vol. 82, p 285.

§ *Ibid.*, vol. 83, p. 85 (1932).

001 of K_2SO_4 is to suppose that when K atoms of the K_2SO_4 structure are laid down in the position described above, they can then act as the foundation for deposition of S_2O_6 groups in the form of the $K_2S_2O_6$ crystal.

VII.—*Conclusion.*

Only the examination of a large number of crystals can show whether the view of adsorption presented here is of general application; but the examples already studied fit in with it so completely that it is worth while examining others to see whether it is a general truth Note that the present view refers only to strong adsorption such as that responsible for modification of habit. For instance, in the growth of a modified crystal it may be—though there is no evidence on this point at present—that adsorption occurs on all faces, but only on certain specific faces with sufficient strength to interfere with growth; on these faces ejection of the impurity is difficult, on all others it is easy; modification of habit is a question of the difficulty of ejection of tightly bound particles.

Adsorption has been studied chiefly in connection with habit modification, *i.e.*, on growing crystal faces. But it is at least as likely to occur when the faces are not growing as when they are; indeed, the formation of oriented overgrowths, which may be regarded simply as adsorbed layers " developed " to visible size, shows that it does occur on stationary faces, and therefore it seems likely that the theory may apply to cases of strong adsorption from solution in general. Whether it is also true (probably with necessary modifications) for strong adsorption of gases on solids is perhaps more doubtful, but there seems to be a growing impression among workers on this subject that the lattice structure of the solid surface is of great importance. This cannot be discussed here.

The present view is from the general standpoint particularly satisfying, for its brings out the close relation between the three modes of intimate association of substances in the crystalline state—adsorption, oriented overgrowth and mixed crystal formation.

The conception on which the present view is based, *i.e.*, the mixed crystal nature of adsorption bodies, and also one of the principal consequences of it, that oriented overgrowth and strong adsorption are due to the same cause, have been anticipated by Gaubert, in 1918 and 1925 respectively, though I was unaware of this when most of the present work was done. Gaubert has not followed up his ideas, nor correlated them; for instance, the 1925 suggestion appears to be made independently of the ideas in the 1918 paper ; and in a later paper (1931) he mentions oriented overgrowths of methyl blue on modified faces of urea oxalate crystals without comment on their significance. The other principle workers on modification of habit at the present time—Buckley in England, and France and his collaborators in America—do not mention Gaubert's ideas. Sloat and Menzies (*loc. cit.*) rejected Gaubert's views, as they failed to obtain oriented overgrowths of urea on octahedral faces of NaCl ; since this has now been achieved, their objections vanish. The present work entirely agrees with Gaubert's views, correlates and amplifies them, and supplies more evidence in support of them.

If the present view is correct, in future the most valuable information on the arrangement of adsorbed particles on crystal faces will be given by oriented overgrowths But it is necessary to know more about the experimental conditions necessary to obtain them.

A study of the known cases of oriented overgrowths should give some idea of the maximum permissible difference of arrangement and of interatomic distances. The work of Barker,* repeated and extended in certain directions by Royer† and Sloat and Menzies (*loc. cit.*), shows that alkali halides of the same lattice structure tolerate a difference in interatomic distance of 15% and 20%, or even 30% for one substance. For substances of different overall lattice structures, no general conclusions on this point have yet been drawn.

The most remarkable thing in Barker's work on the oriented overgrowth of $NaNO_3$ on calcite is that he found that the former grew with its rhombohedral faces in contact not only with the rhombohedral cleavage faces of calcite, but also, much less readily, with natural prism and scalenohedral faces of calcite ; in these positions the only congruence is that of similar edges or zone axes, the faces in contact being quite different This suggests two things First, that in certain examples, probably rare, fairly strong adsorption may occur when there is close correspondence of interatomic distances along one line only. Second, that an adsorbed layer of a substance cannot be arranged like *any* plane in a crystal of the substance, but only like certain specific planes, perhaps the planes which as faces grow fairly slowly Prism and scalenohedral faces are never seen on $NaNO_3$, and therefore have high rates of growth and are unstable planes, and this may be the reason why they do not form adsorbed planes on the corresponding faces of calcite; instead the stable rhombohedral planes of $NaNO_3$ form on the prism and scalenohedral faces of calcite, in spite of the fact that there is correspondence of atomic arrangement only along one line.

---

* 'J. Chem. Soc.,' vol. 89, p. 1120 (1906) ; 'Min. Mag.,' vol. 14, p. 235 (1907) ; vol. 15, p. 42 (1908), 'Z. Kristallog.,' vol. 45, p. 1 (1908).

† 'C. R. Acad. Sci. Paris,' vol. 180, p. 2050 (1925) ; vol. 194, p. 620 (1932).

I wish to thank Messrs. H Emmett and J. L. Matthews for assistance in experimental work I also wish to thank the Directors of Imperial Chemical Industries, Limited, for permission to publish this work, which was carried out in the Winnington Laboratory of I.C.I. (Alkali), Ltd.

Summary.

The properties of crystals usually regarded as being built up by continual adsorption and inclusion of impurity during growth (called here "adsorption bodies") have been compared with those of mixed crystals It is concluded that the two types of crystals differ only in degree. Adsorption bodies may be regarded as unstable mixed crystals.

Applying knowledge on mixed crystal formation in a modified form to adsorption, it is suggested that the condition for strong adsorption is similarity of lattice structure and interatomic distances on specific planes only; the rest of the structures may be quite dissimilar.

This condition is the same as that known to be necessary for oriented overgrowth of different crystals on each other. Another consequence of the ideal is that modification of habit should be reciprocal. A simple explanation of modification of habit is suggested; on the affected faces a mixed crystal is formed, which, being unstable, tends to redissolve, thereby reducing the rate of growth.

Some examples have been studied, and are in complete accord with these conceptions. With NH_4Cl and urea, for instance, the 100 plane of NH_4Cl is similar to 001 of urea, reciprocal modification of habit on these faces occurs, and oriented overgrowths with these faces in contact can be obtained.

Adsorption on crystal faces thus appears to be very closely related to the other modes of intimate association of substances in the crystalline state—oriented overgrowth and mixed crystal formation. An oriented overgrowth is regarded as an adsorbed layer "developed" up to visible size.

If the views stated here are correct, in future the most valuable information on the arrangement of adsorbed particles on crystal faces will be given by oriented overgrowths.

X-Ray Analysis of the Crystal Structure of Durene.

By J. MONTEATH ROBERTSON, M.A., Ph.D.

(*Communicated by Sir William Bragg, O.M., F.R.S.—Received March 21, 1933.*)

1. *Crystal Data, Space Group.*

The crystal structure of durene, 1 . 2 . 4 . 5 . (sym.)-tetramethyl benzene, $C_6H_2(CH_3)_4$, has not previously been examined by the X-ray method, but the class and axial ratios are given by Groth* as monoclinic prismatic

$$a : b : c = 2\cdot4609 : 1 : 1\cdot9975, \qquad \beta = 115^\circ\ 27'$$

with the (100) face most prominently developed. The present work confirms these measurements to within the limits of experimental error (about ½%), but it is found that the axial directions used above define a cell which contains an identical molecule at the centre of the *ac* face. It is therefore necessary to select a new axis in order to define the true unit cell. The relation of Groth's axial directions to the cell now chosen is illustrated in fig. 1, the *b* axis being perpendicular to the paper.

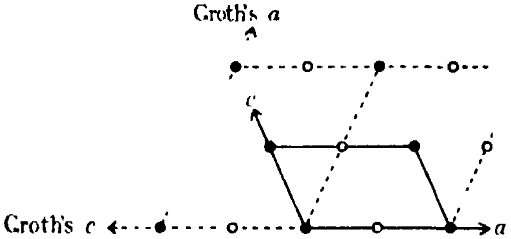


FIG. 1.—Unit cell of durene.

The X-ray work, recorded by means of rotation, oscillation, and moving film photographs, leads to the following values for the axial lengths, etc. :—

$a = 11\cdot57 \pm 0\cdot05$ A.
$b = 5\cdot77 \pm 0\cdot02$ A.
$c = 7\cdot03 \pm 0\cdot05$ A.
$\beta = 113\cdot3^\circ$ $\{h0l\}$ halved when h is odd, (010) halved.

Space group C_{2h}^5 ($P2_1/a$)
Density of crystal $= 1\cdot03$ (Iball)
Number of molecules of $C_{10}H_{14}$ per unit cell $= 2$ ($1\cdot99$)
Volume of unit cell $= 430$ A.^{3}.

---

* 'Chem. Krystallog.,' vol. 4, p. 758 (1917).

Reference to space group literature will show that under these circumstances a centre of symmetry must be assigned to the molecule, and that this is the only symmetry which the molecule utilizes in building the crystal. As far as the space group, then, the structure is similar to certain other symmetrically substituted benzene derivatives, and to the naphthalene-anthracene series. With durene, however, the dimensions of the cell give no clue as to the orientation of the molecules, and the further analysis must depend upon the measurement of the intensities of the X-ray reflections

2 *Experimental Measurement of Intensities.*

The accurate measurement of the intensities of this compound presents considerable difficulties Crystals can be obtained in the form of thin plates or prisms, the large face being the (001) Twinning is usual on this face. The crystals are soft, melting at 80°, and very volatile, a small specimen only lasting about half an hour in the open

It was found that the difficulty of twinning could be overcome by careful crystallization from benzene, when good untwinned specimens were obtained in the form of small prisms. These were cut carefully to the size required (0 4 to 0·1 mg.) with as nearly as possible square sections perpendicular to the zone axis under examination. These small specimens were then sealed off in very thin glass tubes, the uniformity of the glass and its absorption effect on the X-ray beam being tested. Even the thin tubes used were found to stop from 35% to 50% of the copper rays, but, of course, a much smaller fraction of the molybdenum beam.

A complete survey of the relative intensities of the reflections for a number of zones was carried out by the photographic method, employing copper radiation, the procedure being similar to that employed for the anthracene analysis.* Special care had to be taken in estimating the backgrounds owing to extra scattering from the glass tubes. Several short exposures were obtained, however, from a small crystal in the open, by means of one of Dr. Muller's powerful rotating anode X-ray tubes. It was found that these measurements agreed very well with those obtained in the more detailed surveys from the glass protected crystals. As with anthracene, intensities were measured over a range of about 1000 to 1 and special methods were employed to correlate the results over this range.

The general experience is that the smaller crystals lead to higher values for the more powerful reflections The values obtained for these reflections from the smallest available crystals are taken to be correct, without further correction for extinction, and the values are correlated with the measurements of the weaker reflections from the larger crystals used for the extended surveys. An example of this correlation is given in Table I.

---

* 'Proc. Roy. Soc.,' A, vol. 140, p 79 (1933).

Table I.—Integrated intensity, arbitrary units.

hkl	Small crystal unprotected 0·1 mg	Larger crystal under glass 0·4 mg	Combined results
001	643	505	643
002	280	249	280
003	56	58	57
200	357	338	357
600	16	17	17
$20\bar{1}$	499	413	499
$20\bar{3}$	54	53	53
$40\bar{3}$	—	9	9
601	31	29	30
$80\bar{3}$	18	18	18

In other experiments it was found that still higher values were obtained for the (001) and (002), so that the true values for the strong reflections remain a little doubtful. But on the whole quite satisfactory correlations were obtained for the reflections within each zone. In further correlating the separate zones there are, of course, fewer common reflections, and the results are correspondingly more uncertain But again a fairly reasonable agreement was reached in different experiments

So far only the relative intensity measurements have been described. The absolute values of the structure factors were obtained in the following way. Two of the durene crystals upon which an extended series of relative measurements had been made were selected, and a few reflections from each were carefully measured on the ionization spectrometer. These were compared, under exactly the same tube conditions, with two standard crystals of known reflecting power, *i.e.*, two crystals for which the quantity $E\omega/I$ had been accurately measured.* The absorption of the X-ray beam by the glass tubes surrounding the durene crystals was now carefully measured by exploration with fine pin-hole beams of monochromatic copper and molybdenum rays. The correction for the glass tube, and for the absorption in the crystal itself being ascertained, it was now possible to derive the absolute values for the durene reflections. The two crystals employed gave consistent results with copper radiation, but the experiments with molybdenum radiation were somewhat inconclusive. The values used in this paper are taken from the experiments with copper radiation

---

* For the loan of these two standard reflectors I am greatly indebted to Dr B. W Robinson

The two durene crystals for which absolute measurements were thus obtained were now employed as sub-standards to calibrate all the intensity measurements with absolute values With the usual formulæ for the "imperfect" crystal, the values of F were now derived, and these are given in Table II (F measured) Owing to the difficulty of handling these crystals, and the large corrections necessary for the glass containers, etc., the final results are perhaps slightly more uncertain than the anthracene measurements But they are sufficiently accurate to permit a detailed analysis of the crystal structure

Table II.—Measured and Calculated Values of the Structure Factor.

hkl	$\sin\theta$ Cu Kα	F calculated	F measured	*hkl*	$\sin\theta$ Cu Kα	F calculated	F measured.
200	0 145	+27	28	$20\bar{6}$	0 672	– 7	2 5
400	0 290	– 9 5	9	$20\bar{5}$	0 550	+ 2	2 5
600	0 435	–10	12 5	$20\bar{4}$	0 440	– 5 5	8
800	0 580	– 7	10	$20\bar{3}$	0 330	–19	18 5
1000	0 725	+ 2 5	< 3	$20\bar{2}$	0 224	0	5 5
020	0 267	–13	12 5	$20\bar{1}$	0 146	+32 5	34
040	0 534	– 7	4	201	0 222	+13	12 5
060	0 801	+13 5	9	202	0 325	+ 5	5 5
001	0 120	+42 5	38 5	203	0 438	6	5 5
002	0 239	–34 5	37	204	0 551	– 8 5	4
003	0 358	–22 5	20	205	0 668	0	< 2 5
004	0 478	+ 1	4	$40\bar{6}$	0 657	1	< 2 5
005	0 597	– 1	< 3	$40\bar{5}$	0 549	– 9	4 5
006	0 716	– 6 5	3	$40\bar{4}$	0 448	– 3	4
				$40\bar{3}$	0 360	+10	9
011	0 180	+ 3	2 5	$40\bar{2}$	0 291	+ 3	6
012	0 272	– 9	10	$40\bar{1}$	0 266	–14	17
013	0 381	– 6 5	6	401	0 357	– 2 5	< 2
014	0 495	+ 4	< 3	402	0 444	– 1 5	2 5
015	0 610	+ 1	< 3	403	0 545	+ 7	3 5
016	0 725	– 6 5	< 3	404	0 650	+11 5	9
021	0 291	– 6 5	5	405	0 760	+ 3 5	2
022	0 357	– 0 5	3 5	$60\bar{6}$	0 672	+ 9	3 5
023	0 445	– 1 5	3	$60\bar{5}$	0 581	+12 5	3
024	0 545	+ 4	2	$60\bar{4}$	0 500	– 2	2
025	0 653	+ 5	4 5	$60\bar{3}$	0 439	–19 5	12
026	0 761	+ 1	< 3 5	$60\bar{2}$	0 403	– 8	9
031	0 417	+ 0 5	3	$60\bar{1}$	0 402	+ 8 5	2
032	0 465	+ 1	< 2 5	601	0 496	–24	18 5
033	0 537	0	< 2 5	602	0 576	– 2	< 2 5
034	0 621	– 2	< 2 5	603	0 665	+18 5	11·5
035	0 717	1 5	< 2 5	604	0 765	+12	5·5
041	0 546	– 6 5	5	$80\bar{6}$	0 718	+17	9·5
042	0 583	– 4	< 3	$80\bar{5}$	0 645	+12	6 5
043	0 642	+ 1	< 3	$80\bar{4}$	0 585	– 9	9
044	0 715	+ 6	2	$80\bar{3}$	0 545	–14	14·5
051	0 678	– 0 5	< 2 5	$80\bar{2}$	0 530	– 3 5	< 2·5
052	0 710	+ 1 5	< 3·5	$80\bar{1}$	0 544	– 2 5	< 2·5
				801	0 638	– 2	3
				802	0 711	+ 2 5	2·5

Table II—(continued).

hkl.	sin θ Cu Kα	F calculated	F measured	*hkl*	sin θ Cu Kα	F calculated	F measured.
$100\bar{5}$	0·731	+1	< 2·5	550	0·758	+0·5	< 4·5
$100\bar{4}$	0·690	+3·5	< 3	610	0·455	−3	4·5
$100\bar{3}$	0·668	0	2·5	620	0·510	4	3
$100\bar{2}$	0·665	−5·5	3	630	0·589	0·5	< 3·5
$100\bar{1}$	0·686	2·5	< 4	710	0·525	−5	3·5
				720	0·574	−0·5	< 3·5
110	0·152	+29	27	810	0·595	4	3
120	0·276	−6·5	13	910	0·666	5·5	2
130	0·407	+3	< 2				
140	0·539	+6	2	$11\bar{4}$	0·470	9	5
150	0·670	+2	< 3	$11\bar{3}$	0·360	−7	4·5
210	0·196	+49	56·5	$11\bar{2}$	0·256	−5	5
220	0·303	+12	14	$11\bar{1}$	0·174	+20	19
230	0·426	+1	3	111	0·210	+11·5	10
240	0·553	+1·5	< 3·5	112	0·305	9·5	9
250	0·682	11·5	8	113	0·415	−7·5	7·5
310	0·256	+23	30·5	$12\bar{1}$	0·290	+2	< 3·5
320	0·345	+17·5	14·5	$13\bar{1}$	0·415	−5·5	7
330	0·455	−8	5	$21\bar{2}$	0·260	11·5	12·5
340	0·575	−6	6	$22\bar{3}$	0·422	−5	< 5
350	0·700	+7	5	$22\bar{2}$	0·347	−9·5	10·5
410	0·319	+8·5	10·5	$22\bar{1}$	0·304	4·5	< 3
420	0·393	+12	5·5	221	0·347	−8·5	11
430	0·492	+3·5	< 3·5	222	0·421	17	15
440	0·606	+6·5	3·5	$23\bar{2}$	0·459	3	< 5
450	0·725	−4	2	$31\bar{3}$	0·362	+17	15
510	0·386	−4	< 3	$32\bar{3}$	0·430	+7·5	6
520	0·450	+13	9·5	$33\bar{3}$	0·523	+2·5	< 5
530	0·539	+11	8	$33\bar{1}$	0·449	8·5	5·5
540	0·644	−6·5	3·5	$41\bar{4}$	0·468	−8	5

3. *Analysis of the Structure.*

Even with the assumption that only the carbon atoms have any appreciable scattering power for X-rays, the durene structure still contains 15 independent parameters. The problem is therefore much too complex to be undertaken without some further guiding principles. The chemical structure states the relative positions of the atoms in the molecule, but the exact form and dimensions of the ring and side groups are unknown. Now, the results of the anthracene investigation have shown that even in a complex structure where the molecule does not coincide with any crystal plane, the aromatic rings are regular plane hexagon structures, to quite a high degree of accuracy. It seems a reasonable assumption, then, to take a plane model of the durene structure, based upon the anthracene dimensions, as shown in fig. 2.

The carbon to carbon distance in the ring is 1·41 A., and the CH_3 groups are placed at the greater distance of about 1·56 A. from the benzene ring, but a small variation of the latter distance is quite unimportant at the present stage of the analysis.

The problem now reduces to the much simpler one of finding the orientation of this structure in the crystal. The success with which the measured structure factors can be explained by means of this model must be the justification for using it. The finer details of the structure, *e.g.*, the exact position and scattering power of the CH_3 groups, can only be settled by a more detailed Fourier analysis. Unfortunately, any such Fourier analysis must be preceded by a careful trial and error analysis, in order to decide the phase constants of each term.

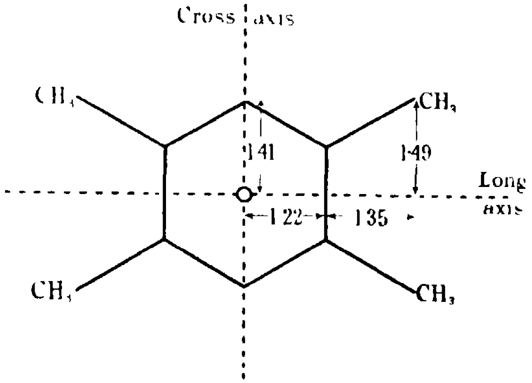


FIG. 2. —Durene model.

The durene crystal resembles the naphthalene-anthracene type by its tabular or flaky form, the (001) being by far the most important face. It can at once be shown, however, that there is no possibility of the molecules lying in this plane. If they did, the structure factor for the (001) would have a value of about 106* and the higher orders would fall off normally in intensity. The measured value of F(001) is actually only 39, with almost normal decrease to the (003), when the value drops abruptly almost to zero.

Analogy with the naphthalene-anthracene structures, and the pseudo-orthorhombic nature of the lattice (compare fig. 1) suggests the possibility of the long axis of the molecule being nearly normal to the (001) plane. But this possibility can be entirely ruled out by a simple calculation which gives the values shown in Table III for the structure factors for this orientation.

---

* In these calculations the graphite-anthracene atomic *f*-curve for carbon is employed. The formulæ for calculating the geometrical structure factors for this space group have already been given ('Proc. Roy. Soc.,' A, vol. 125, p. 542 (1929)).

Table III

hkl.	F calculated Long axis normal to (001).	F calculated Cross axis normal to (001).	F measured
001	3	43	39
002	2	38	37
003	2	24	20
004	4	5	4
005	27	2	<3

However, when the cross axis of the molecule is placed normal to the (001) plane, the agreement between the calculated and the observed values for these reflections is very close, as shown by the above figures.

These examples are sufficient to illustrate the method by which the correct orientation of the molecule can be found. Other sets of reflections can be employed to find how the planes of the molecules are inclined to each other In general the structure is found roughly from the low index reflections, and is refined by a careful study of the higher index reflections, which are more sensitive to small movements.

The best agreements have been obtained for the following orientation. χ, ψ and ω are the angles which the long axis of the molecule makes with the *a*, *b* and *c′* (perpendicular to *a* and *b*) crystal axes, χ', ψ' and ω' are the corresponding angles for the cross axis of the molecule (compare fig. 2). Any three of these quantities can be treated as independent parameters. Then

$$\chi = 47^\circ \qquad \chi' = 97^\circ$$
$$\psi = 43^\circ \qquad \psi' = 90^\circ$$
$$\omega = 85^\circ \qquad \omega' = 7^\circ$$

The structure factors calculated for this position are compared with the measured structure factors in Table II.

4. *Discussion of Results.*

It will be seen that the average agreement between the calculated and observed values is fairly good, and is at least sufficient to determine the sign, or phase constant of all but a few of the weaker reflections beyond any reasonable doubt. Trials of different positions seem to show that while movements of up to 3° have not much effect, yet a movement of 5° in any direction will destroy many of the agreements found. There remains the possibility that some distortion of the model used, *e.g.*, in regard to the position of the CH_3

groups, combined with a movement in some direction, might improve the agreements between the calculated and the observed values. Owing, however, to the good results obtained for all the stronger reflections, it seems clear that such alterations must be of a small order of magnitude. It is hoped to obtain a clearer picture of the fine detail of the structure by means of a double Fourier analysis which is now being undertaken. It is important to emphasize, however, that such a Fourier analysis must depend in its main outlines upon the analysis now given.

The outstanding features of the structure are already clear, and the result is somewhat surprising. The durene molecule has a practically flat or disc-like form as with anthracene and hexamethyl benzene, but these discs are arranged almost perpendicularly to each other, instead of being parallel as they are in hexamethyl benzene.* The arrangement is best seen by projecting the structure normally on to the (001) plane, as shown in fig. 3. Owing to the fact that the molecules are tilted only about 7° away from the normal to this plane ($\omega' = 7°$) the skeleton molecules of fig. 2 project almost into straight lines in this diagram.

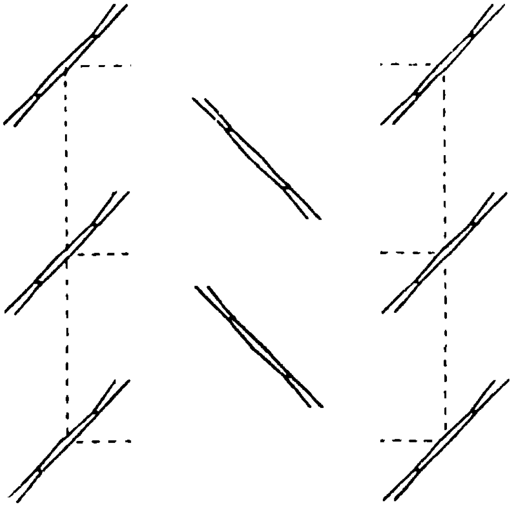


FIG. 3.—Projection on the (001).

The closest sideways distance of approach between atoms in adjacent molecules is 3·71 A., a figure which may be subject to a small variation when the structure is refined. But it is remarkably similar to the corresponding figure for anthracene of 3·77 A. The gap between the ends of the molecules along the *c* axis is of the order of 4·2 A., a figure again very similar to the anthracene result (4·06).

---

* Lonsdale, 'Proc. Roy. Soc.,' A, vol. 123, p. 499 (1929).

With regard to the other physical properties of the crystal, I am indebted to Mr J. D. Bernal for an optical examination. His results are consistent with the molecular orientation arrived at in this analysis, the refractive indices showing that there is no possibility of the molecules being parallel.

In conclusion I wish to thank Sir William Bragg, O M., F.R.S., and the Managers of the Royal Institution for their interest in this work.

5. *Summary.*

A quantitative X-ray examination of the crystal structure of durene, 1 . 2 . 4 . 5-tetramethyl benzene, is described. The space group is C_{2h}^5 ($P2_1/a$), and there are two molecules in the unit cell. The structure is thus similar to the naphthalene-anthracene type, but the cell dimensions give no clue as to the orientation of the molecules, which in this case must be deduced directly from the intensity measurements.

A technique is described for dealing with the small volatile crystals. Integrated intensity measurements have been made, and the absolute values of the structure factors obtained for about 100 reflections.

In the analysis of the structure it is shown that to a first approximation at least, the molecules are flat or disc-like, as in anthracene and hexamethyl benzene. The precise orientation of the molecules has been obtained from a study of the intensities and is given on p 600. It is found that the planes of molecules in neighbouring rows are arranged almost perpendicularly to each other, instead of being parallel, as might be expected. The closest distance of approach between atoms in adjacent molecules is about 3·7 A., very similar to the distance found for anthracene (3·8 A.).

The Spectra of the Halogen Molecules. Part I.—*Iodine*

By W. E. Curtis, D.Sc., and S. F. Evans, M.Sc., Armstrong College, Newcastle-upon-Tyne.

(Communicated by T. H. Havelock, F.R.S.—Received March 23, 1933.)

[Plate 9.]

Introductory.

The term "halogen molecule" is intended to include the interhalide compounds as well as the elementary molecules. The existence of three of the ten possible molecules of this type, namely, IF, BrF, and ClF, has not yet been established, and we have not attempted any spectroscopic search for them. For F_2 the data are meagre, only an emission spectrum being known,* and we have therefore confined our attention to the other six molecules, namely, Cl_2, Br_2, I_2, BrCl, ICl, and IBr. The spectra of these, with the exception of BrCl, have been extensively investigated, and partially interpreted. There are certain features, however, notably the continua and diffuse bands observed in emission, which have not yet been satisfactorily explained, and we have set out to survey the existing data systematically and, where necessary, to obtain new data with the object of accounting for these features.

The problem of obtaining satisfactory data of continua and diffuse bands is definitely different from that involved in the case of line and ordinary band spectra. Not only is it more difficult to make accurate visual settings for the purpose of wave-length determinations, but it becomes of much more importance to record the intensity distribution, which is often very characteristic and undoubtedly significant. For these reasons we have thought it essential to obtain and reproduce microphotometer records in all experiments, and to determine wave-lengths solely from these records. Quite low dispersion has been used, and appears very suitable for a general survey, although a more detailed investigation using higher dispersion may well prove profitable later. It seemed important to us to get strictly comparable data in all observations, and it has therefore been necessary to redetermine some of the existing data, which although numerous are seldom complete or homogeneous. A number of entirely new data are also presented. Some consideration must first be given to the known band systems of these molecules, since they provide evidence relating to the location and characteristics of molecular levels, some at least of which may be expected to be concerned in the emission of the continua and diffuse bands.

---

* Gale and Monk, 'Phys. Rev.,' vol. 29, p. 211 (1927); vol. 33, p. 114 (1929).

There are such strong general similarities between the spectra of all the halogen molecules that it is safe to assume that the same general interpretation will be applicable to them all The data are by no means equally extensive; for example, the banded absorption in the visible region which is shown by the vapours of I_2, Br_2, Cl_2, ICl and IBr has not been observed at all in the case of F_2, presumably because the absorption coefficient is so small that an excessively long column would be required Again, the IBr molecule is so unstable that it is difficult to obtain its spectrum free from I_2 bands, whilst the still greater instability of BrCl in the vapour phase has so far prevented the detection of any characteristic absorption, although an emission spectrum undoubtedly exists, as will be shown.

On account of the general correspondence between these spectra it should evidently be profitable to consider them together and especially to seek to establish between them as detailed a correspondence as possible, so that tentative interpretations may be tested and accidental numerical coincidences more readily exposed This is what we have attempted to do, but to simplify the discussion it is desirable at first to consider only a single typical molecule, that of I_2, for which the data are much more complete than for any other of these molecules

The Spectrum of I_2.

In view of the complexity of the phenomena and the number of investigations concerned it is convenient to treat the absorption, fluorescence, and emission spectra separately. Their chief features are indicated diagrammatically in fig. 1.

(*a*) *Absorption.*—The best known feature of the absorption spectrum is the band system (B) extending from the red to the green and merging at the convergence point into its associated continuum (C). There is also a much weaker band system (A) in the near infra-red, with a continuum (not shown) having a maximum about 7300 A. Another set of bands (D) occurs in the ultra-violet, between 2750 and 1700 A., but it is not yet certain whether these all belong to one system. Between 2750 and 2090 A. Pringsheim and Rosen* have measured some 300 bands in iodine vapour at a pressure of about one atmosphere and at temperatures up to 800° C. Kimura and Miyanishi† used a lower pressure and a temperatures of about 120° C., and observed 157 bands in the region 2145 to 1948 A. Sponer and Watson* measured bands overlapping this system and extending down to about 1770 A. These appear to represent an extension of Kimura and Miyanishi's system but are relatively more intense. The same observers also discovered a group (E) of very strong bands in the region 1770 to 1670 A., and some further absorption (F) between 1600 and 1500 A., but experimental difficulties prevented the detection of structure, if any. The investigation of the absorption spectrum in the region of still shorter wave-lengths has not yet been attempted.

---

* 'Z. Physik,' vol. 50, p. 1 (1928).

† 'Sci. Pap. Inst Phys. Chem. Res. Tokyo,' vol. 10, p. 33 (1929).

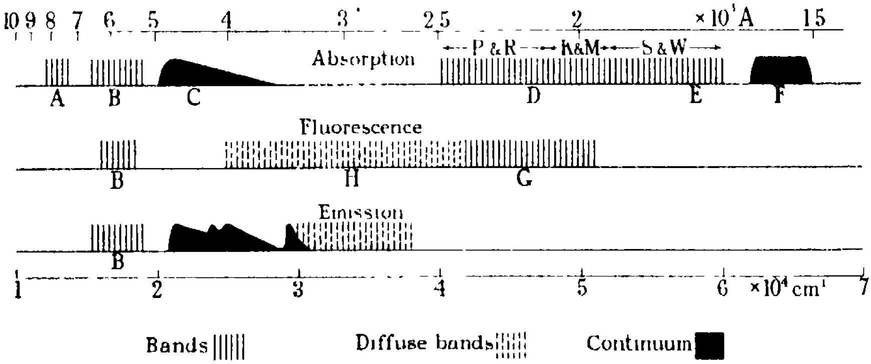


FIG. 1.—Spectra of molecular iodine.

(*b*) *Fluorescence.*—It has long been known that iodine vapour at low pressures can be excited to fluorescence by absorption of Hg 5461 or other line in this region. The resulting doublets (or groups if a wide exciting line is used) have been extensively studied, originally by Wood, and completely interpreted in terms of the current theory of band spectra. Indeed, they provide one of the most direct and striking applications of this theory. They are indicated under B in fig. 1. There is a further extensive fluorescent spectrum which can be excited by radiation of 1900 A. or thereabouts. For a short distance above the exciting wave-length this spectrum consists of a regular series of sharp lines (G) closely resembling Wood's series in the visible region and admitting of a similar interpretation. But after about the 35th member they merge into a series of diffuse bands (H) which are at first fairly regularly spaced but which later show no obvious regularity at all. These extend to about 4000 A.†

---

* 'Z. Physik,' vol. 56, p. 184 (1929).

† McLennan, 'Proc. Roy. Soc.,' A, vol. 91, p. 23 (1914); Oldenburg, 'Z. Physik,' vol. 18, p. 1 (1923).

The three continua at 3460, 4300, and 4800 A. which are prominent features of the emission spectrum can also be obtained by ultra-violet irradiation in the presence of a foreign gas, nitrogen for instance, at considerable pressures, but Oldenberg concludes that this is not genuine fluorescence but photochemiluminescence

(c) *Emission* —The emission spectrum varies very greatly with the pressure and current density. If both are sufficiently high the atomic spark spectrum may readily be obtained, with, of course, the arc spectrum, but free from molecular spectrum. In the present work we are only concerned with the latter, and conditions have accordingly been chosen to give the molecular spectrum as free as possible from the atomic. The high frequency discharge (about 10^7 oscillations per second) was employed throughout, with external electrodes, and the pressure of the iodine vapour was about 0·1 mm or less, being controlled by the temperature of a side limb. In the green-red region a set of bands (B) may be obtained which are the counterpart of the ordinary absorption bands. They are not identical, however, since the intensity distribution is dependent in the case of emission on the electrical conditions and in the case of absorption by the temperature of the vapour. It is noteworthy that the continuum which is associated with the absorption bands is not also observed in the spectrum of a discharge tube, although it has been obtained as part of the recombination spectrum of heated iodine vapour.*

The most characteristic and best known features of the emission spectrum are the three continua already mentioned, which are usually designated by their approximate positions 4800, 4300, and 3460 A They differ markedly in appearance and vary in relative intensity with the conditions. 4800 is sharp on the red side, falls off in intensity slowly to the violet and probably extends well into the ultra-violet. According to Bhatnagar, Shrivastava, Mathur, and Sharma† it terminates definitely at 2130 A , but we have not been able to confirm this 4300 is probably also sharp on the red side, but as it is overlaid by 4800 it is not possible to compare the two edges satisfactorily. It shows traces of structure on one of our plates taken in the second order of a 21-foot grating. We have found that there is also a maximum at about 4020 A. which is particularly prominent at low pressures. 3460 is distinctly different in character, being approximately symmetrical and comparatively narrow (about 100 A.). Further, it is much more persistent than 4800, which is only prominent over a limited range of pressure, both in emission and fluorescence. It is obtainable in fluorescence in the presence of a large excess of foreign gas (*e g.*, N_2 at atmospheric pressure), when an extensive system of diffuse bands is apparently associated with it.*

---

* Kondratjew and Leipunsky, 'Trans. Faraday Soc.,' vol. 25, p. 736 (1929).

† 'Phil Mag.,' vol. 5, p. 1226 (1928).

There is finally a considerable number of diffuse bands, more or less regularly arranged, between 2700 and 3300 A These appear to be in the main identical with those observed in fluorescence, except that the long wave members of the latter system do not occur in emission (under the conditions we have used) and the short wave members are masked, if present, by overlying continuum. The only published measurements of these bands in emission are those of Bhatnagar and his co workers (*loc cit*), which refer to the edges and not to the intensity maxima. Since the latter are very well defined on our microphotometer records, and since, too, the published values of Oldenberg and McLennan for the fluorescent bands are not in very satisfactory agreement we have thought it worth while to make fresh wave-length determinations The results are given in Table I, and microphotometer traces of spectra at low and higher pressures are reproduced in Plate 9 The wave-lengths given are probably correct to within an Angstrom or two, except for broad or unsymmetrical maxima.

Interpretation of the Absorption Spectrum of I_2

The absorption spectrum is relatively the simplest of the three, since one level, that corresponding to the normal state of the molecule, is common to all the systems. This is therefore the natural starting point for an investigation of the whole spectrum The data have been summarized above. In the visible and near infra-red regions their interpretation is established, but the position with regard to the ultra-violet is at present somewhat unsatisfactory. Fig. 2 shows the approximate form of the potential energy-nuclear separation curves as derived from Morse functions based on the experimental data The normal state of the molecule is $^1\Sigma_g^+$ and the upper level of the visible bands is one component of $^3\Pi_u$, designated by Mulliken† O^+ and dissociating into atoms in the $^2P_{1/2}$ and $^2P_{3/2}$ states. The infra-red system‡ arises from a transition from the ground state to another of the components of $^3\Pi_u$ which has the same dissociation limit as the normal state of the molecule There are a number of other theoretically possible molecular states based on the same two dissociation limits, but so far there is no experimental evidence of their existence, unless a fragmentary band recorded by Nakamura,* which does not appear to belong to either of the known systems, should prove to be associated with one of them.† There is also a third mode of dissociation possible, namely $^2P_{1/2} + {}^2P_{1/2}$, but there is at present no experimental evidence relating to this. Some of the states in question are doubtless of the repulsive type, but these would give rise to continua if they combined with the ground state, and the only continua observed up to the present are those associated with the visible and infra-red systems. The latter of these is very weak. The former is stronger and more extensive, and is probably composite, including separate continua due to $v'' = 0$, 1, and so on, but no very detailed examination of its structure appears to have been made

---

* Oldenberg, 'Z. Physik,' vol. 25, p 136 (1924).

† 'Phys. Rev.,' vol. 36, p. 1440 (1930).

‡ Pringsheim and Rosen, *loc. cit.*; Brown, 'Phys. Rev.,' vol. 38, p. 1187 (1931).

The chief present problem in connection with the absorption spectrum is that of the interpretation of the ultra-violet band system or systems Pringsheim and Rosen, who studied the long-wave end of this, concluded that it arose from a transition from a lower state B ($\omega \simeq 90$ or 180) lying slightly above the normal, A ($\omega = 214$), to an upper state D ($\omega \simeq 70$) Sponer and Watson have proposed an alternative arrangement in which these bands are linked up to the main system farther to the ultra-violet, thus rendering it unnecessary to postulate the additional low level B. There are two difficulties in accepting this suggestion. One is that the Pringsheim-Rosen bands do not appear to join on to the others very satisfactorily There is only a long single progression forming a sort of bridge between them, and looking more like a $v'' = 0$ progression than one in the middle of a system. On the other hand, it must be remembered that the data are not a homogeneous set, but are due to different observers working under widely different conditions of pressure and temperature Since in addition the accuracy of the measurements is low, discrepancies of as much as 10 cm.$^{-1}$ appearing in the mean intervals, it is hardly possible to reach a definite conclusion at present More accurate measurements are very desirable, and preparations to this end are now in hand in this laboratory.

The second difficulty in Sponer and Watson's interpretation is that very high values of v'' (up to 28) must be assumed to occur, and it seems at first unlikely that there could be a sufficient concentration of such molecules to give appreciable absorption. We have, however, made some calculations bearing on this point and find that at 200° C. (Mecke's maximum temperature), the concentration of molecules having $v'' = 9$ (the highest value observed by him) is approximately $1 \cdot 4 \times 10^{-3}$, whereas at 800° C. (Pringsheim and Rosen's maximum temperature), the concentration for $v'' = 28$ is $7 \cdot 9 \times 10^{-5}$. The relative pressures and lengths of path used are somewhat uncertain, and dissociation at the higher temperature has been neglected, but it does not appear by any means impossible that absorption could be observed up to $v'' = 28$. In any case the data are especially uncertain for the highest v'' values, and it may be that the last few progressions are spurious. On the whole there does not seem to be sufficient justification for Pringsheim and Rosen's assumption of an additional low state B, although it is not improbable that more than one upper state is involved.

---

* 'Mem. Coll. Sci. Kyoto Imp. Univ.,' A, vol. 9, p 315 (1926).

† But Pringsheim and Rosen reported that they were unable to obtain this band, and Mr. O. Darbyshire, working in this laboratory, has been equally unsuccessful.

The interpretation of the far ultra-violet system of Sponer and Watson is fairly clear. The intensity distribution is of a type characteristic of a transition which is accompanied by a marked loosening of binding. The first progression ($v'' = 0$) is very long, and extrapolates to a convergence in the neighbourhood of 60 200 cm.$^{-1}$ * The vibrational level intervals in the upper state are known from the analysis, but the v' values are arbitrary. The rotational analysis has not been attempted, and the nuclear separation r is therefore not obtainable directly. But a rough estimate of its value may be obtained by making use of Morse's approximate relationship $\omega_0 r_0^3$ = constant. For the two states concerned in the visible absorption bands this product has the values 4020 and 3500. Assuming a mean value of 3700, r_0 for the state in question comes out to about 3·5 A. The product $\omega_0 r_0^2$, which is often nearly constant for different electronic states of the same molecule, has considerably different values in this case, namely, 1510 and 1160.

The remaining quantities necessary for the evaluation of an approximate curve relating potential energy and nuclear separation, namely, ω_0 and $\omega_0 x$, have been estimated from the data of Sponer and Watson. The values upon which the curve shown in fig. 2 is based are as follows: $\omega_0 = 90$ cm.$^{-1}$, $\omega_0 x = 0 \cdot 11$ cm.$^{-1}$, D = 15,400 cm.$^{-1}$, $r_0 = 3 \cdot 5$ A. The usual Morse formula $U = D\,(e^{-a(r-r_0)} - 1)^2$ has been employed.

The curve is admittedly somewhat hypothetical, but so far as it can be tested it appears to fit the facts reasonably well. It accounts, for example, for the general distribution of intensity in the band system and for its change

---

* This is the sum of the wave-number of the 0, 0 band (44,800) and of the dissociation potential D (1·9 volt), both as estimated by Sponer and Watson. They do not give the convergence wave-number explicitly, and it is evidently subject to considerable uncertainty.

with temperature, as well as for some of the fluorescence phenomena to be discussed later It exhibits several unexpected features, however ; the dissociation work (15,400) is considerably larger than that of the normal state

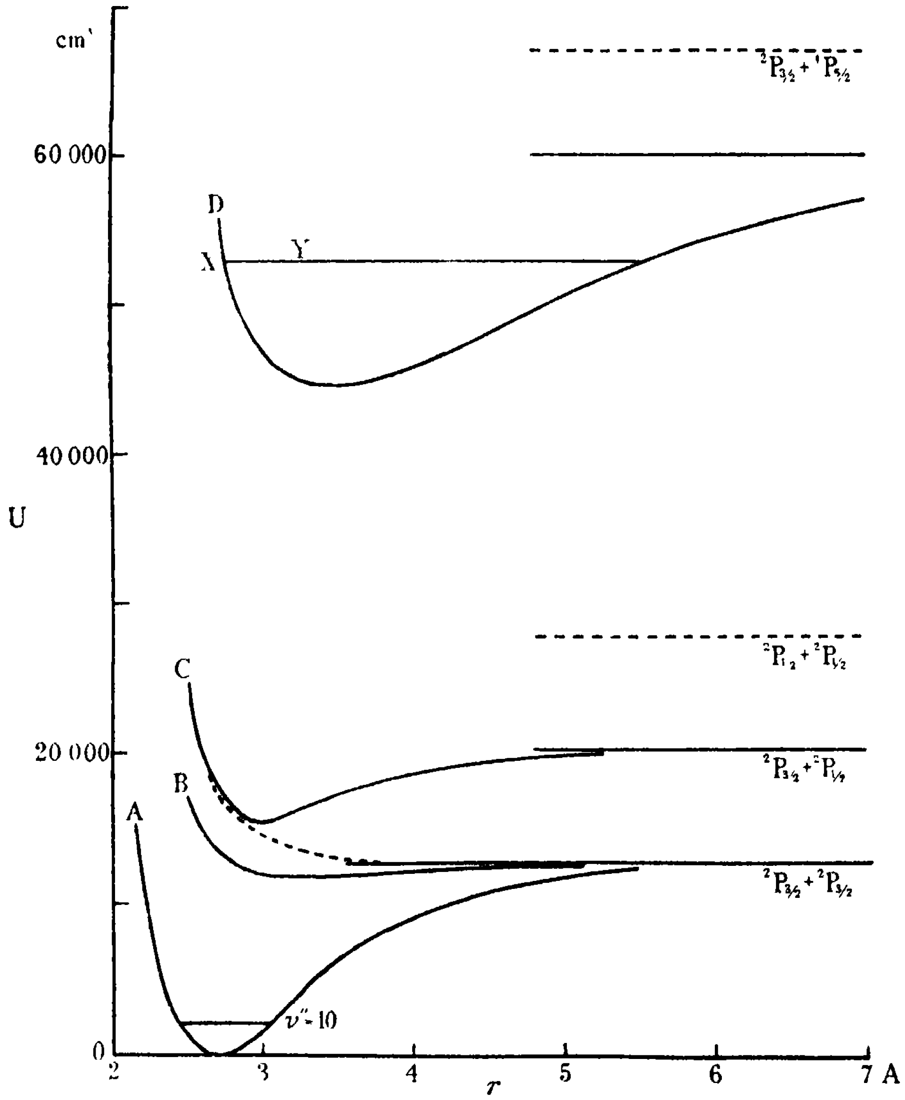


Fig. 2.—Potential energy curves for I$_2$.

(12,400), which is quite exceptional for a molecule of this type, although it is characteristic of certain unstable molecules, such as Hg$_2$. The curve is also abnormally broad, owing to the extremely small value of *x*, which in its turn denotes a close approximation of the nuclear vibrations to simple harmonic type Both of these features would suggest that the excited state in question differs in some fundamental way from the others, and this is borne out by a consideration of the dissociation limit

This is approximately 60,000 cm $^{-1}$ above the ground state of the molecule, which corresponds to a total atomic excitation of about 47,600 cm. $^{-1}$, assuming, as appears to be well established, that the ground state dissociates into two unexcited ($^2P_{3/2}$) atoms. Now the first atomic state above the normal is $^2P_{1/2}$ at 7600, and the upper state of the visible absorption bands has its convergence at just this distance above the dissociation limit of the ground state, thus dissociating into one $^2P_{3/2}$ and one $^2P_{1/2}$ atom. The next limit, corresponding to dissociation into two $^2P_{1/2}$ atoms, would lie at 15,200, but no molecular states converging to such a level are known, although Kondratjew and Leipunsky* have reported an I_2 excitation potential at 3·8 volts which appears to refer to this level. The next possibility is $^2P_{3/2} + {}^4P_{5/2}$, a total excitation of 54,632, or 7000 cm. $^{-1}$ above the estimated convergence of the ultra-violet system. There appear to be three alternative explanations of this discrepancy which may be advanced, namely:

(1) That Sponer and Watson's value for the convergence, 60,200, may be some 7000 too low. An error of this magnitude would appear to be out of the question, however.

(2) That the ground state may dissociate into one $^2P_{1/2}$ and one $^2P_{3/2}$ atom instead of into two $^2P_{3/2}$ atoms as hitherto assumed. This is at first sight rather a tempting hypothesis, since it would just about account for the magnitude of the discrepancy. There is no direct evidence bearing on the question, since the chemical determination of the heat of dissociation of the normal I_2 molecule gives no information as to the nature of the products. But there are two distinct lines of indirect evidence which support the accepted view. One is derived from a correlation of atomic and molecular states,† and the other from the interdependence of the heats of dissociation of I_2, Cl_2, and ICl. From the spectroscopic data for these molecules, and assuming the products of dissociation to be known in each case, one can calculate the heat of dissociation for normal ICl.‡ The only way in which agreement with the chemical determination can be secured is to suppose that normal I_2 gives two normal atoms on dissociation, in accordance with the accepted view.

---

* 'Z. Physik,' vol 44, p. 708 (1933).

† Mulliken, 'Phys. Rev.,' vol. 36, p. 1440 (1930).

‡ Cordes, 'Z. Physik,' vol. 74, p. 40 (1932).

(3) That the dissociation products are not excited atoms. The only alternative would appear to be ions, I^+ and I^-, in either their normal or some excited state.* The energy of I^+ relative to I has been determined by the electron impact method by Foote and Mohler† and from an analysis of the atomic iodine spectrum by S. F. Evans.‡ The results are respectively 10·5 and 10·44 volts, a much better agreement than could be expected in view of the long and uncertain extrapolation upon which the latter is dependent.§ Taking the convergence wave number for the level in question as 60,200, and subtracting the energy of dissociation of I_2 into normal atoms, namely 12,400, we get on the present hypothesis the value 47,800 for the energy of $I^+ + I^-$ relative to normal atoms. Adopting 10·5 volts (85,100 cm.$^{-1}$) as the energy of I^+, the electron affinity is 47,800 − 85,100 = − 37,300 cm.$^{-1}$.

Values for the electron affinity have been obtained by two other methods. One of these is based on calculations of grating energies of crystals, the most recent and reliable value being that of Mayer and Helmholz.‖ They consider that their result, 26,090 cm.$^{-1}$ (74·2 k. cal. per mol.) should be accurate to within about 2%. The other method is due to Knipping,¶ and is based on his experimental determination of the work required to dissociate HI into $H^+ + I^-$. As some of the data he employed have since been revised it is repeated below.

He considers the energy changes involved in the cyclic set of changes

$$HI \rightarrow H^+ + I^- \rightarrow H^+ + I + \varepsilon \rightarrow H + I \rightarrow \tfrac{1}{2}H_2 + \tfrac{1}{2}I_2 \rightarrow HI$$

Introducing the energy changes, all supposed positive, we have

$$\begin{aligned} &HI + V_{HI} = H^+ + I^- \\ &H^+ + I^- + E = H^+ + I + \varepsilon \\ &H^+ + I + \varepsilon = H + I + V_H \\ &H + I = \tfrac{1}{2}H_2 + \tfrac{1}{2}I_2 + \tfrac{1}{2}D_H + \tfrac{1}{2}D_I \\ &\tfrac{1}{2}H_2 + \tfrac{1}{2}I_2 = HI + Q, \end{aligned}$$

---

* This suggestion has been advanced on other grounds by several other workers, namely, Oldenberg ('Z. Physik,' vol. 18, p. 1 (1923)), Sponer and Watson, ('Z. Physik,' vol. 56, p. 195 (1929)) and Mulliken ('Phys. Rev.,' vol. 36, p. 1447 (1930)).

† 'Phys. Rev.,' vol. 21, p. 382 (1923).

‡ 'Proc. Roy. Soc.,' A, vol. 133, p. 417 (1931).

§ The value 10·548 recently published by Deb ('Proc. Roy. Soc.,' A, vol. 139, p. 380 (1933)) is based on the assumption that the "3460" continuum is emitted as a result of the capture of an electron by an ionized I atom. The energy required to produce such an ion from the normal molecule is about 97,000 cm.$^{-1}$, whereas it is possible to excite the where V_{HI} is the work of dissociation of HI into $H^+ + I^-$, E is the electron affinity of I, *i e.*, the work required to detach the extra electron from I^-, V_H is the ionization potential of H, D_H and D_I are the heats of dissociation of H_2 and I_2 into normal atoms, Q is the heat of formation of HI from H_2 and I_2 as determined chemically and ε represents a free electron.

Adding and equating the two sides we get

$$V_{HI} + E = V_H + \tfrac{1}{2}D_H + \tfrac{1}{2}D_I + Q.$$

The values of the various quantities are as follows --

V_{HI} = 12·7 volts (Knipping, *loc cit*, corrected by Gerlach and Gromann) = 103,000 cm.$^{-1}$

V_H = 109,700 (Rydberg constant)

D_H = 35,800 (spectroscopic value given by Professor O W Richardson in a private communication)

D_I = 12,400 (spectroscopic value)

Q = 450 (1290 cals per mol, see *e g* Saha and Srivastava, ' Heat," p. 463).

Hence E = 31,250 cm.$^{-1}$. This is 5000 higher than the value deduced from grating energies, which as stated above is regarded as being within 2% of the true value. The only one of the above quantities which can be seriously in error is V_{HI}, and it therefore seems desirable that it should be redetermined.

If we accept the Mayer-Helmholz value for the electron affinity as the most reliable at present available, we have to account for the discrepancy of some 11,000 cm $^{-1}$ which results from the adoption of hypothesis (3). We might assume the products of dissociation to be excited, with energies 85,100 + X for I^+ and − 26,090 + Y for I^-. We should then have 47,800 − 85,100 − X = − 26,090 + Y, and therefore X + Y = −11,210. Since both X and Y are essentially positive the difficulty is only accentuated by such an assumption. It is possible, but distinctly improbable, that the above value of the atomic ionization potential is in error by so much as 11,000 cm. ^{1}. On the whole, it seems best to regard the question of the dissociation products of the level in question as still an open one.

---

emission of " 3460 " by absorption of radiation in the neighbourhood of 53,000. It follows that the suggested interpretation is untenable, and that no weight can be given to the value of the ionization potential so derived.

|| ' Z. Physik,' vol. 75, p. 19 (1932).

¶ ' Z. Physik,' vol. 7, p. 328 (1921).

There is, however, one other point of interest in connection with these bands. It might appear that the question as to whether the dissociation products are ions or excited atoms could be answered by observing whether radiation of suitable frequency produces conductivity in the vapour. Ludlam and West* have, in fact, found such an effect using radiation from an aluminium spark, but the effective wave-length was uncertain, and no very definite conclusions could be drawn. If there is a continuum associated with this system it should be about 60,000 cm.$^{-1}$, *i.e.*, 1670 A. or beyond. A continuum in this region (F of fig. 1) has, in fact, been observed by Sponer and Watson.

There remains to be considered the banded absorption of still shorter wave-length discovered by Sponer and Watson. This comprises some six very intense bands in the neighbourhood of 1750 A., together with some weaker bands extending down to 1666 A. The former group shows intervals approximating to 213 cm.$^{-1}$, the separation of the ground state vibrational levels, and has been interpreted by Sponer and Watson as a transition to an upper level having approximately the same nuclear separation and vibration frequency as normal I_2. That is to say, they regard each band as a sequence of heads arising from the same Δv, thus accounting for their remarkably high intensity. This view receives some support from observations of the effect of raising the temperature to 140° C.; this does not affect the number of bands but merely broadens the middle members. This would correspond to a strengthening of those components of a sequence which arose from higher v'' values. At the same time there should be a transfer of intensity within the system towards longer wave-lengths, but this is not recorded. Another difficulty in the sequence interpretation is that the bands are apparently symmetrical, whereas an unresolved sequence should certainly be sharper on one side or the other, according to whether ω' or ω'' is the larger. The dispersion used was small, however, and may have been insufficient to show such an effect if present. The chief argument in favour of the sequence interpretation seems to be the absence of an alternative. The only possible one would be to attribute the bands to a transition from the normal state to an upper level of very low stability, corresponding to a nearly flat potential energy curve. This would give diffuse bands having separations of the observed order, but their intensities should fall off steadily from the high frequency end, which would correspond to absorption from the lowest and most abundant vibrational state, whereas, in fact, they show a maximum near the middle. Also the effect of increasing temperature would be to bring up further members at the long-wave end, but Sponer and Watson state definitely that this does not occur Further experimental work is necessary before the question can be settled, but as the evidence stands at present one must regard the sequence interpretation as the more acceptable. This necessitates an upper level of similar characteristics to the ground state, with its lowest vibration level about 57,000 cm^{-1} above that of the latter On the assumption of a similar work of dissociation this is also the total excitation energy of the dissociation products. These might be atoms in the normal ($5p\,{}^2P_{3/2}$) and $6s\,{}^4P_{5/2}$ states, having a total energy of 54,600, or the excited atom might be in the next higher state, $6s\,{}^4P_{3/2}$, giving a total energy of 56,100

---

* 'Proc. Roy. Soc., Edin.,' vol. 45, p. 34 (1924).

The remaining absorption bands in this region, which lie mainly on the short-wave side of the above group, have not been analysed We have detected some possible regularities, however, as set out in the Table I. For comparison the main group of Sponer and Watson is also included. In the absence of any intensity estimates the groupings can only be regarded as tentative, but it may be worth recording them so that they may be tested when better data become available.

Table I.

Series I.		Series II		Series III.		Series IV	
57,517		58,386		58,600		60,006	
	179		224		264		238
338		162		336		59,768	
	195		262		277		222
143		57,900		059		546	
	198		305		268		218
56,945		595		57,791		328	
	225		(347)		274		223
720		*(248)		†517		105	
	219		(390)				
501		858					

* Interpolated, not observed, but may be under 57,143, which is very broad
† Is also the first member of Series I.
Series I is Sponer and Watson's "group of six" referred to above

Interpretation of the Fluorescence Spectrum.

The fluorescent series associated with the visible band system has been extensively studied and need not be considered here. We shall confine our attention to the ultra-violet system. It has been excited by various radiations between 1800 and 2100 A., and extends from the exciting wave-length up to about 4000 A. The high frequency part up to 2150 A. is a typical molecular fluorescence series converging towards the red and showing the intervals characteristic of the lower state. It may be assumed to originate in absorption of one or more of the lines of the ultra-violet system, raising the molecule to the excited state D at about the level marked X, fig. 2. The series of 35 sharp fluorescent lines correspond to transitions from the latter back to the lowest 35 levels of the normal state. The remaindor of the fluorescence spectrum, consisting of some 70 more or less diffuse bands, is not so readily explained.

Kuhn* has suggested an interpretation of those, 25 in number, lying at the low frequency end, *i e.*, between 3169 and 4007 A. From a graph of the intervals, reproduced in fig. 3, he concludes that the bands form one series, due to the transition from an upper firmly bound state ($\omega \sim 220$ cm $^{-1}$) to a lower loosely bound or unstable state. There are several difficulties in his explanation, however. In the first place, although Kuhn remarks that Sponer and Watson have found an excited state with this vibration frequency, this cannot be the upper state in question, since it lies too high ($\sim$ 57,000) above the ground level to be capable of excitation by the frequencies used (none above 54,000). Further, the necessity for postulating such a state does not arise if the criticism of Kuhn's graph which follows is sound An examination of the diffuse bands recorded by McLennan between 31,000 and 38,000 cm.$^{-1}$ strongly suggests that they form one series, with certain members missing. The first three gaps are each near strong Hg lines (3125, 3025, 2967) which would almost certainly have obscured any bands in their immediate neighbourhood under McLennan's experimental conditions. The remainder cannot be so explained, but it is probably significant that the next four gaps occur between successive bands; that is to say, that alternate members appear to be suppressed (in the region 34,500 to 36,000). By interpolating these, together with several others between 36,000 and 38,000 a fairly satisfactory complete series is obtainable. The resulting intervals are plotted in fig. 3 and indicate that at least two distinct series exist, one between 4007 and 3394, converging rapidly towards the violet and the other below 3394 converging slowly in the same direction. This conclusion is supported by the intensities, which show a definite discontinuity at the same point and also by the fact that we have obtained the latter series very readily in emission without any trace of the former. Incidentally it may be noted that the convergence to the violet negatives the possibility of either series being of the pure fluorescent type, since any emission series originating in a single upper level must converge to the red. The intervals plotted in fig. 3 are taken from our own data derived from the emission spectrum and given in column 4 of Table II. McLennan's fluorescence results are given in column 2 and although showing large systematic differences from the corresponding emission bands evidently refer to the same series.

---

* 'Z. Physik,' vol 63, p. 475 (1930).

Of these two series, the long wave one shows a very rapid change of intervals, which no longer approach a limit (~ 220 cm $^{-1}$) as in Kuhn's interpretation. The larger intervals occurring are much above any vibration frequency to be expected for the iodine molecule, but may possibly arise as Kuhn suggests from a combination of an upper firmly bound state with a lower loosely bound or unstable state. The interval graph is very similar to that obtained in the case of certain Hg_2 bands discussed by Kuhn, except that in the present case the limiting value of the intervals, which should correspond to the vibration frequency of the upper state, is not approached, and this cannot therefore be identified. We have not found it possible to construct a plausible lower state potential energy curve from the data on the assumption that the state D is

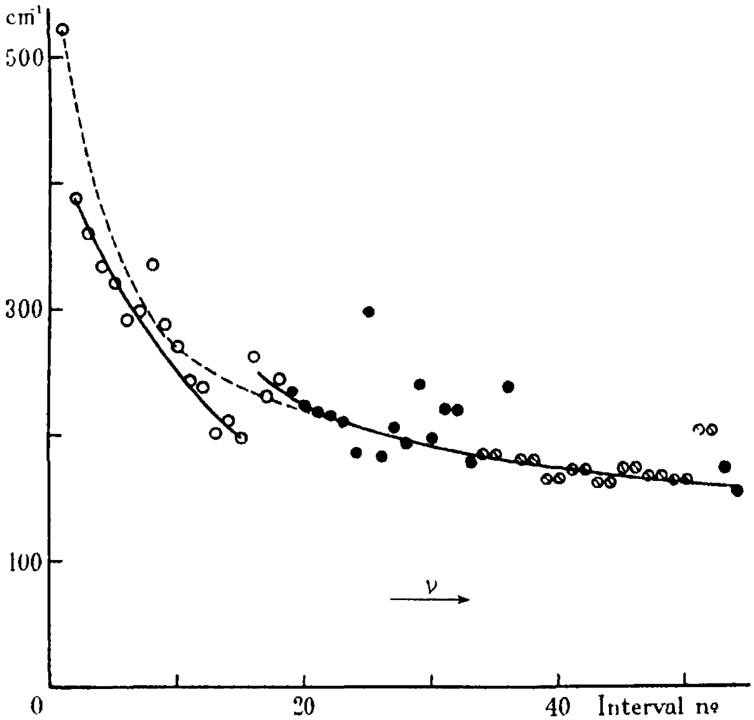


FIG 3.—Intervals from ultra-violet diffuse bands. ⊙ Intervals from Oldenberg's data ; ● intervals from present data, ⦸ half-intervals from present data, - - - - Kuhn's curve.

the upper state associated with these bands. It is more probable that the latter is arrived at by a transition from D, which would account for the comparatively large number of upper vibrational states involved.

The above observations will apply equally well to the second or short wave series. Qualitatively both series might be accounted for as arising from a transition from a stable upper level to different portions of the same unstable lower level. The data at present available do not permit of the location of either of these levels. It is only possible to say that the lower level is probably associated with the limit $^2P_{3/2} + ^2P_{3/2}$, and that the upper cannot be far below the level D. These conclusions follow immediately by subtracting the maximum frequency of the system ($\sim$ 37,000 cm $^{-1}$) from the excitation frequency in fluorescence ($\sim$ 53,000) which gives 16,000 as the maximum possible excitation of the final state. If there is an intervening transition the latter would be still lower. But an unstable lower level must lie above the fundamental limit $^2P_{3/2} + ^2P_{3/2}$, *i e.*, 12,500 cm $^{-1}$ and this leaves at most 3500 for the preliminary transition. There is one group of diffuse bands in the neighbourhood of 31,000 cm.$^{-1}$ (3220 A), that is to say, at the beginning of the short wave series, which is so prominent as to call for special mention. There are some seven or eight of them, with intervals ranging from 205 to 245 cm.$^{-1}$ and intensities rising to a maximum near the middle. The irregularities in the intervals, although considerable, probably do not exceed the errors of measurement, and the intensities run so smoothly that one is bound to conclude that the group is a real one, with some simple physical connection between the members. The spacing and intensity distribution are very similar to those of the group near 1700 A. observed by Sponer and Watson in absorption, but we have not found it possible to trace any precise correspondence between them nor is it likely that such should exist, since the absorption group arises from the excitation of a much higher level than that concerned in the emission of the fluorescent bands.

It has been suggested above that the greater part of the fluorescent band system may be derived from an upper state or states lying below the state D and excited by a preliminary transition from it. A detailed consideration of the properties of the latter state lends support to this view. According to the Franck-Condon theory an electronic transition in a molecule has little or no direct and immediate effect on the position or motion of the nuclei. The two conditions taken together would suffice to determine the frequency capable of being emitted by a molecule at any phase of the nuclear vibration. Although the principle is not to be interpreted too precisely, it is of great use in indicating the most probable transitions. In the present case we may apply it to the level X of the state D in order to predict the nature of the fluorescent radiation. Since the vibration period (10^{-12} second) is small compared with the average life of the excited state ($\sim 10^{-8}$ second ?) we may assume that molecules in all phases of vibration will be found, with increasing probability towards the turning points.

The kinetic energy is represented by the vertical height of the level above the potential energy curve. So long as this is less than the corresponding distance (*i.e.*, for the same nuclear separation) in the ground state a transition to the latter may take place, but if it is greater continuous emission resulting in dissociation would occur. That is to say, only molecules represented by the points between X and Y can give discontinuous fluorescence by transitions to the ground state. Any other state with the same dissociation limit (12,400) is bound to be less stable than the latter, and the above restriction will operate to an even greater degree It follows that the extensive observed system of fluorescent bands cannot originate directly from the level D There must be an intermediate transition to a third level, or perhaps more than one Such levels would not combine with the ground state and would not therefore show in absorption. The fact that the fluorescence spectrum is practically independent of the exciting wave-length, so long as this lies between 1850 and 2000 A., also strongly suggests that a secondary transition must intervene between those of excitation and fluorescence.

Interpretation of the Emission Spectrum.

Mention has already been made of the principal features of the emission spectrum. Except for the bands in the less refrangible region of the visible spectrum, none of them have as yet received a really satisfactory explanation. Considering first the relatively broad continua, we find that there is general agreement, on experimental grounds, in assigning those near 4800 and 4300 to the iodine molecule.* We have found, using high dispersion, distinct signs of a complex band structure in the case of 4300 excited by a low pressure discharge, and Cario and Oldenberg† have reported a regular structure in the case of 4800 excited in a narrow capillary. It thus appears probable that neither of these is a genuine continuum. There is no such evidence relating to the emission maximum which we have observed near 4020 A., but there is no reason to doubt that it is molecular in origin.

---

* *E.g.*, Finkelnburg, 'Phys. Z.,' vol. 31, p. 1 (1930); Meeke, "Handbuch d. Physik" (Geiger and Scheel), vol. 21, p. 640 (1929).

† 'Z. Physik,' vol. 31, p. 914 (1925).

Concerning the maximum usually referred to as "3460"* there has been a good deal of controversy. It differs markedly in behaviour from the others. It is much more persistent and is especially prominent at higher pressures, when it appears to break up into a regular series of diffuse bands (Cario and Oldenberg, *loc. cit.*) It is unique, moreover, in that it remains very prominent in highly dissociated iodine vapour, and attempts have therefore been made to ascribe it to the iodine atom. The occurrence of structure seems to preclude this, however, and at present no satisfactory way out of the dilemma is apparent. Gerlach and Gromann† attributed it to the capture of an electron by the iodine atom, but the resulting value (28,760 cm.$^{-1}$) for the electron affinity differs considerably from that of Mayer and Helmholz (26,090) derived from grating energies. Oldenberg suggested that it might result from the recombination of I^+ and I^-, thus accounting for the simultaneous atomic and molecular characteristics as arising from the initial and final states respectively. We cannot find any numerical support for this, since using the values quoted above for the ionization potential and electron affinity, the excitation of the resulting iodine molecule then comes out to be 42,200 cm.$^{-1}$, which is a rather unlikely value as will be seen by reference to fig 2. There is also considerable doubt as to whether such a process as $I^+ + I^- = h\nu + I_2$ is possible No such ionic recombination with emission of radiation has as yet been established, although a number of molecular states are known which dissociate adiabatically into ions. It is true that Terenin and Popow‡ have reported several observations of photodissociation into ions but they do not appear to have shown that this takes place in one act. It is more probable that the absorption of light results in excitation to an unstable level, which then dissociates into ions.

The most promising clue in connection with 3460 is a recent observation of Skorko§ that it can be obtained at high temperatures in absorption and then shows a coarse vibration structure. By following this up it may be possible to establish the position of the upper level and hence determine its nature.

The wave-lengths and intensities of the system of diffuse bands, as determined from our microphotometer traces, are given in Table I, together with the

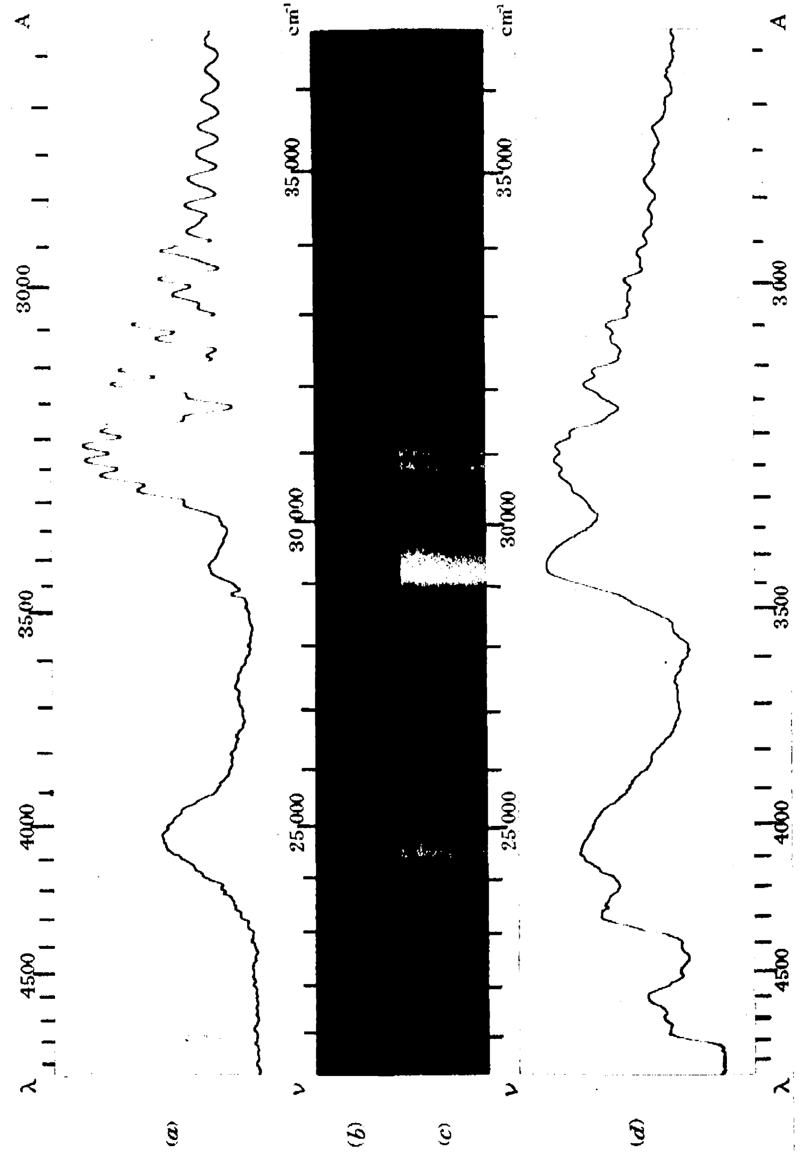
---

* This wave-length refers to the long-wave edge. Although this is comparatively sharp no accurate position can be assigned to it, and we have accordingly determined the wave-length of the intensity maximum (3414 A., see Table II).

† 'Z. Physik,' vol. 18, p. 239 (1923).

‡ 'Z. Physik,' vol. 75, p. 338 (1932).

§ 'Nature,' vol. 131, p. 366 (1933).


A
3000
3500
4000
4500
λ
cm⁻¹
35 000
30 000
25 000
ν
(a)
(b)
(c)
(d)

corresponding data of previous observers. The plates were taken on a Hilger small quartz spectrograph having an average dispersion of 46 A. per millimetre in this region, and the traces themselves are on a scale of about 4½ A. per millimetre. The pressure of iodine was of the order of 0·001 mm. The spectrum and microphotometer record are reproduced as (*b*) and (*a*), Plate 9. The effect of raising the pressure to 0·16 mm. (corresponding to atmospheric temperature) will be seen from (*c*) and (*d*) The relative intensity of " 3460 " is much greater, and the continua at 4800 and 4300 are prominent. Otherwise the spectrum is essentially the same, except that the diffuse bands are less conspicuous on account of the increased intensity of the continuous background. It is impossible to say whether this is an extension of 3460 or a separate continuum, but the former seems quite probable.

The first two columns refer to low pressure fluorescence, and the third to the Tesla discharge as studied by Bhatnagar, Shrivastava, Mathur, and Sharma. The latter workers have measured both edges of the bands; the tabulated values are the means of these.

The intensities attached to our values are measured directly from the microphotometer traces and thus have no absolute significance. But they have the advantage of being objective, and do represent with some accuracy the order of intensities in any particular region. The intensities (in brackets) given by other observers are presumably estimated visually in the usual manner.

It is evident from the table that the fluorescent and emission spectra are essentially identical in this region. At the long-wave end the agreement is complete, except for systematic wave-length differences, whilst at the short wave end there are some slight discrepancies There is good general agreement between our results and those of Bhatnagar and co-workers. They list a few additional weak bands which are, in fact, present on our microphotometer traces, although not suitable for measurement. But we cannot confirm their conclusion that all the above bands form an ordinary system, with " 3460 " as the (0, 0) transition. In the first place, the relative enhancement of " 3460 " by pressure is so pronounced as to put any such simple relationship out of the question. Further, the majority of the diffuse bands have a remarkably symmetrical intensity distribution, whereas an ordinary band is necessarily unsymmetrical in this respect. Until they have been examined under higher dispersion too much insistence should not be laid on this evidence. But a study of the proposed $v'v''$ scheme on the basis of our data does not provide any support for it, since the combination differences vary irregularly by 100 cm.$^{-1}$ and more from the average values. Such discrepancies might be per-

Table II.—Diffuse Bands in Low Pressure Spectrum of Iodine.

Oldenberg.	McLennan	Bhatnagar and co-workers.	Curtis and Evans.			
λ and Int.	λ	λ and Int.	λ.	Int.	ν.	
		3400	3414·3	3·6	29,280	
3311 (7)	3315	3307 (2)	3306·4	3·8	30,236	
3285 (9)	3290	3277 (5)	3280·9	6·2	30,471	235
3261 (10)	3268	3251 (8)	3256·9	8·2	30,695	224
3237 (10)	3245	3228 (8)	3233·8	8·7	30,914	219
3212 (10)	3220	3206 (8)	3211·4	8·8	31,130	216
3191 (9)	3195	3183 (8)	3189·8	7·9	31,341	211
3169 (6)	3175	3164 (3)	3171·0	3·7	31,527	186
		3142 (3)	3141·3	3·5	31,825	298
		3119 (8)	3123·3	7·5	32,008	183
	3107	3101 (7)	3103·3	6·9	32,214	206
	3090	3083 (1)	3084·8	2·7	32,408	194
	3065	3059 (6)	3062·1	4·6	32,648	240
	3047	3040 (7)	3043·6	5·9	32,846	198
		3022 (1)	3023·3	1·5	33,067	221
	3009	3002 (6)	3003·3	3·5	33,287	220
	2993	2985 (7)	2987·2	4·4	33,466	179
	2960	2970 (1)				
	2946	2951 (7)	2954·6	3·5	33,836	370
	2930	2938 (2)	2933·9	2·8	34,074	238
	2915	2920 (0)				
	2900	2908 (2)	2903·1	3·0	34,436	362
	2883	2872 (5)	2875·4	2·9	34,768	332
	2853	2846 (3)	2847·0	2·7	35,114	346
	2825	2820 (3)	2820·9	2·5	35,439	325
	2799	2794 (0)	2793·3	2·3	35,789	350
	2774	2765 (1)	2767·2	2·2	36,127	338
	2760	2748 (1)				
	2737	2739 (1)	2742·0	2·0	36,459	332
	2715	2728 (1)	2711·5	2·6	36,869	410
	2697		2698·6	2·6	37,045	176
	2685		2687·2	2·9	37,202	157

missible so long as the edges of the bands were measured (as by Bhatnagar and co-workers), since they could be attributed to blending, but the increased accuracy of our results and the symmetry of the maxima renders such an explanation untenable. Apart from this the ω' value proposed ($\sim$ 710) is so high as to be exceedingly improbable. The fact that the intensity maxima in the sub-groups follow a fairly smooth curve, and also that all the bands vary together with pressure, suggests that they may belong to one system. On the other hand, the grouping and intensities do not resemble those of any known system. Broadly speaking, three regions may be distinguished. At the long wave end, between 3350 and 3150, is the very prominent group already mentioned, with quite regular intervals and intensity distribution. The intervals converge from 235 to 211 cm.$^{-1}$ towards the ultra-violet, and the intensities increase in the same direction, reaching a maximum for the penultimate band, namely, 3211 A., and then falling slightly for 3190. The next band, at 3171, shows an abrupt decrease of intensity and interval, and probably does not belong to this group, although usually included in it.

At the short wave end, from 2950 to 2700 A., we have another set of bands, showing considerable regularity of intensities and intervals, but the latter are much larger than in the previous group, ranging around 340 cm.$^{-1}$, and decrease more slowly in the same direction. In the intermediate region there are three groups each containing three bands, with intervals ranging from 180 to 240 and little apparent regularity as regards intensity. A fundamental difficulty in connection with the short wave set is the wide spacing, 340 cm.$^{-1}$. Since the vibrational frequency in the ground state is only 213 cm.$^{-1}$ it is very unlikely that this should correspond to an actual vibrational frequency of the excited molecule. There are two ways in which intervals much greater than the actual vibration frequencies may occur. Firstly, by the combination of a stable level with one of low stability, provided that relative positions of the potential energy curves are favourable. In such a combination, however, the intervals must show a rapid variation if they differ considerably from those characteristic of the stable state,* and this explanation is therefore excluded by their comparative constancy in the present observation. Secondly, it is possible to obtain a pseudo-interval as the result of a fortuitous relationship between the vibration frequencies of two combining states. If one, for example, is approximately two-thirds of the other, coincidences will occur between every third band of the former progression and every second of the latter. In the present observation, the observed maxima with a spacing of 340 cm.$^{-1}$ might be due to the superposition of two progressions having intervals of about 170 and 133 cm.$^{-1}$. If this is the true explanation, some confirmation should be obtainable from microphotometer traces of much higher dispersion, since the coincidence will not be exact over more than a very narrow range.

---

* See Kuhn, 'Z. Physik,' vol. 63, p. 461 (1930)

Returning now to consideration of the prominent group about 3250 A., the main question is whether the approximate agreement of the observed intervals (235 to 211 cm.$^{-1}$) with those of the ground state ($\omega_0'' = 213$ cm.$^{-1}$) is significant or accidental. It cannot actually be that of the ground state, or these bands would appear in absorption. If the excitation to the upper level takes place by electron impact all trace of ω'' necessarily disappears. If the excitation takes place by absorption of radiation, as, for example, of the atomic resonance line 1844, which is very strong in the discharge, there might appear to be a possibility of the population of the upper state being to some extent dependent on ω'', since a group of lower levels with this spacing will have been separately excited by the same frequency. Since the rotational energy has to be taken into account as well, however, the upper levels excited will not show the ω'' spacing. We conclude therefore that the intervals of this group probably have nothing to do with those of the ground state, and this is supported by the fact, which has already been mentioned, that the intervals show a rapid variation, such that if they tend to a constant value it must be well below 200 cm.$^{-1}$.

We have recently found that a very similar set of diffuse bands can be produced in bromine at very low pressures; the study of these, which is now in progress, may very well throw some light on the question at issue.

We desire to express our indebtedness to the Department of Scientific and Industrial Research for a grant which made this investigation possible.

Summary.

The chief features of the I_2 spectrum which have not yet received an explanation are described and discussed. In the absorption spectrum the chief problem is the nature of the excited states associated with the various systems. Of these the best determined is the upper level of Sponer and Watson's ultra-violet system. A potential energy curve is given for this, and is found to possess some unusual characteristics. Consideration of these, together with the apparent impossibility of accounting otherwise for the dissociation energy of the products, leads to the conclusion that the latter are probably ions.

The resulting value for the electron affinity of the iodine atom is not, however, in satisfactory agreement with the latest determinations from grating energies.

Some possible regularities in the far ultra-violet absorption bands are proposed, but more precise data, including intensity estimates, are much to be desired in this region.

The fluorescence system of diffuse bands due to ultra-violet excitation is next considered, and it is shown that Kuhn's conclusions require modification. The bands allocated by him to one series actually fall into two, one of which converges rather rapidly with increasing frequency whilst the other converges more slowly. It is remarkable that the bands still farther to the ultra-violet may be regarded as an extension of the latter series provided that alternate members are supposed to be missing. This suggestion is admittedly without theoretical support, but the alternative difficulty of accounting for the wide spacing of the bands ($\sim$ 340 cm.^{-1}) is almost equally serious. It is concluded that the emission of the diffuse bands cannot immediately follow the absorption of the exciting radiation, but that a second transition intervenes.

New data derived from microphotometer records of the diffuse emission bands are presented. These do not provide any support for the attempt made by Bhatnagar, Shrivastava, Mathur, and Sharma to fit the bands into an ordinary $v'v''$ scheme.

Experiments on Molecular Scattering in Gases. I.—*The Method of Crossed Molecular Beams.*

By Ronald G. J. Fraser, Imperial Chemical Industries, Limited, and L. F. Broadway, Clare College, Cambridge.

(Communicated by T. M. Lowry, F.R.S.—Received March 28, 1933.)

The development of the technique of molecular rays has made possible a direct experimental study of collision phenomena in gases. Hitherto our knowledge of molecular fields has been derived indirectly; mainly from experiments on gaseous viscosity and diffusion, which demand inevitably a statistical interpretation. On the other hand, direct measurement of the angular distribution of the molecules scattered from a beam in traversing a gas at low pressure is now practicable; and it seemed worth while to develop a method of investigating intermolecular collisions which should make use of the new technique.

The experiments described in this paper are a first step in this direction. Ideally one would wish to study the mutual scattering of two similar or dissimilar beams of nearly homogeneous velocity; but it was considered advisable, in orienting experiments such as the present, to aim at a compromise. In the first place, therefore, attention has been confined to the scattering of one beam only, which we may call the primary beam. In the second place, the choice of substances for the other beam, the scatterer, has been confined to those of which the molecules are very heavy and sluggish compared with the molecules of the primary beam; in such case, the scattering of the primary beam is sensibly symmetrical about its own direction. As constituents of the primary beam we are using in the first instance the alkali metals, mainly because there already exists for them in the surface ionization gauge an extremely sensitive detector.*

Outline of Method.

The experimental arrangement is shown schematically in fig. 1. A beam of alkali metal atoms of circular cross-section is directed to pass across and immediately above the circular orifice of an oven, from which a heavy vapour is effusing molecularly to be condensed on a cooled hood above the oven. The density of the vapour stream is adjusted so that the alkali metal atoms suffer single collisions in passing through the apex of the stream. The scattered atoms are received on a circular ring of tungsten wire, which forms the positive electrode of a surface ionization gauge. The gauge is movable along the axis of the primary beam; in this way the angle of collection of the scattered atoms can be varied at will over a prescribed range. For a given angle of scattering the number of atoms falling on unit area of the detector per second is inversely as the square of its distance from the centre of scattering; hence a comparatively large number of scattered atoms is received at the larger angles. This enables measurements to be made with approximately equal accuracy over the whole range of a steeply falling angular distribution curve. This advantageous arrangement is, of course, only possible if the scatterer is a heavy, low velocity molecule; otherwise the circularly symmetric distribution of the scattered atoms about the axis of the primary beam is lost.

---

* The surface ionization gauge was discovered by Langmuir and Kingdon (' Proc. Roy. Soc.,' A, vol. 107, p. 61 (1925)), and adapted to the measurement of the intensity of alkali metal beams by Taylor (' Z. Physik,' vol. 57, p. 242 (1929)). It is described in some detail in Fraser, "Molecular Rays," Cambridge (1931), p. 43 ff.

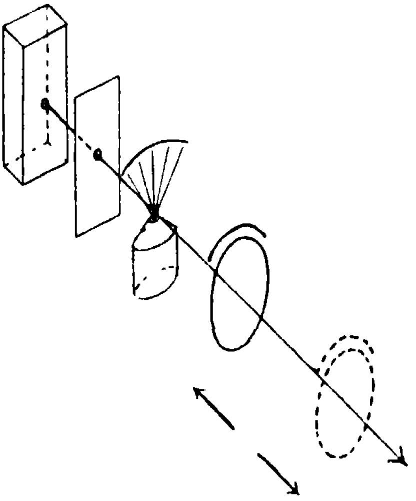


Fig. 1.

The Apparatus.

The apparatus is illustrated in fig 2. It consists essentially of a brass tube, divided into two chambers, the oven chamber and the observation chamber, by the liner (L), which carries the beam system of oven (O_1) (on a dovetailed slide (D*t*)) and image aperture (I), and which can be removed after each experiment for cleaning. Each chamber communicates via a mercury trap with a Leybold steel diffusion pump type B; pressures were read on a McLeod gauge. The oven (O_2) containing the scatterer is hung on two constantan girders from the Dewar vessel (D_3) containing liquid air, which is provided with a copper hood (H) to condense the molecules effusing from O_2 The detecting system, consisting of the tungsten filament (F) and surrounding cylinder (C), is carried on the nut (N) of a lead-screw (L*s*), and can be traversed along the axis of the primary beam by means of the screw driver (Sc); two spiral springs (S*p*) of phosphor bronze ribbon form extensible leads to the detector. The filament (F) is protected from grease vapour by the large copper shield (Cs), which is attached to the Dewar vessel (D_4) containing liquid air. It may be mentioned here that the lead screw is lubricated with low pressure apiezon grease.

The dimensions are as follows: oven aperture and image aperture, each 0·5 mm. diameter; distance apart, 5·7 cm. Distance of orifice of O_2 from image aperture, 1·2 cm.; diameter of orifice, 2 mm. Distance of ring from scattering centre at forward extreme of traverse, 3·2 cm.; distance at backward extreme of traverse, 13·6 cm. Diameter of ring, 1·0 cm. Range of angle covered, 2° to 9·3°.

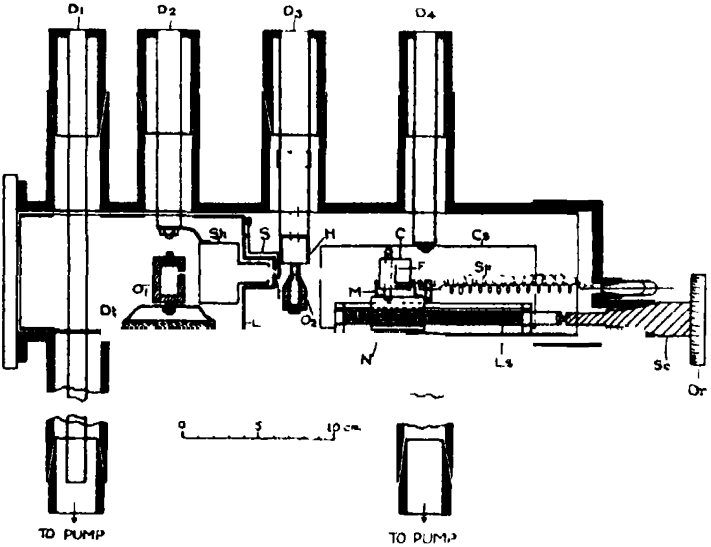


FIG. 2.

Adjustment.—The adjustments necessary are three in number: (1) the beam is adjusted parallel to the direction of motion of the detector system along its ways; (2) the circular filament is adjusted symmetrically about the beam; (3) the oven (O_2) is adjusted so that the primary beam passes about 1 mm. above the orifice under the temperature conditions obtaining during an experiment. These adjustments were carried out optically.

The Primary Beam.—Direct comparison of the scattered intensity at a given angle with that of the primary beam would have introduced serious additional complication into the design of the detecting system. Since a knowledge of the relative amounts of scattering at different angles is sufficient for our present purpose, attention was rather directed to maintaining the primary beam as steady as possible over periods of several hours. To this end particular attention was paid to the design of the oven O_1, and to the control of its temperature during an experiment; to the provision of a large clean surface of molten alkali metal in the oven, and to the vacuum conditions in the oven chamber.

The oven was of high-speed tool steel, copper plated and gilded to ensure small radiation loss and a uniform temperature distribution. It was heated by radiation from two tungsten spirals carried each in a vertical hole bored in the body of the oven. The temperature of the oven was measured by means of an iron-constantan thermocouple and millivoltmeter. When the desired temperature was reached, the e.m.f. of the thermocouple was balanced out on a potentiometer used in conjunction with a sensitive galvanometer; 1 cm. deflection corresponded to a temperature variation of 0·25°. The oven temperature could readily be maintained constant to well within this limit by a hand-controlled resistance in the oven heating circuit.

The internal diameter of the oven (10 mm.) was large compared with that of the oven orifice (0·5 mm.); the metal was degassed and filtered *in vacuo* before introduction into the oven, which could then be rapidly inserted on its dovetailed slide into position in the apparatus. Thus a clean molten surface of area large compared with that of the orifice was ensured.

A large liquid air cooled surface was provided in the oven chamber by the shield (Sh) attached to the Dewar (D_2) in order to ensure the efficient condensation of the alien atoms of the alkali metals.

As a result of these precautions, it was possible to maintain the intensity of the primary beam constant to within ½% for several hours.

The beam could be admitted to the observation chamber at will by means of the magnetically operated shutter (S).

The Scattering Oven.—The oven O_2 was of copper-cadmium alloy, the heating system being similar to that of the oven O_1; with liquid air in the Dewar D_3, a temperature range of from $-100°$ C. to $+100°$ C is easily covered, and any chosen temperature within this range could be maintained extremely constant for hours.

The orifice was given the form of a short circular canal, 4 mm. long, diameter 2 mm., in order to concentrate the effusing molecules about a vertical direction.*

The Detector System.—The circular filament was of spectroscopically pure tungsten wire, 0·1 mm. in diameter, for which we have to thank the General

---

* *Cf.* Clausing, 'Z. Physik,' vol. 66, p. 471 (1930).

Electric Company. It was run at a temperature of 1400° K.; one end was connected to a phosphor bronze spring, the other was earthed to the brass casing of the apparatus. To secure approximately uniform temperature throughout the ring, the free ends were bent at right angles to the plane of the ring, and were carried back a distance of about 15 mm. before being attached to the spring lead and carriage respectively. A light mask (M) was carried on the front of the detector carriage to prevent scattered atoms striking these horizontal portions of the filament wire, which would otherwise present greater effective areas of collection at forward positions of the detector.

Before each experiment, the filament was flashed for one minute at 2700° K.; in order to avoid damage to the spring leads through overheating, the insulated end of the filament was brought into contact with a heavy lead while flashing. With sodium, the filament was oxygenated, after first flashing, by admitting air from the fore-vacuum, with the filament at a temperature of 1900° K.; the filament temperature was reduced to 1400° K. before re-establishing high vacuum.*

The cylinder, at a potential of about -20 volts, was connected to earth via the second phosphor bronze spring and a Leeds and Northrup moving coil galvanometer, having a current sensitivity of 2×10^{-11} amps./mm. at 1 metre.

With this arrangement, an intensity of only 3×10^{10} atoms/cm.^{2} sec., that is about a ten-thousandth of the intensity of the primary beam, could be measured correct to 2%.

In order to make possible a rapid traverse of the detector system, the lead screw is furnished with a three start thread of 4·5 mm. pitch. The thread having been calibrated, the position of the ring is determined by counting the number of complete turns of the screw driver from the zero position by means of a cyclometer; fractions of a turn are read off on a graduated drum, Dr, fig. 2. Positions of the detector can then at once be translated into terms of the angle of collection knowing the diameter of the ring and its distance from the scattering centre at the zero position.

Experimental Results.

For the present orienting experiments, *trans*-di-iodoethylene, IHC = CHI, was chosen as scatterer. (1) The molecular weight is 280 as against 23 and 39 for sodium and potassium respectively, (2) the mean molecular velocity is $1·5 \times 10^4$ cm./sec., as against $8·6 \times 10^4$ cm./sec. for sodium and $6·2 \times 10^4$

---

* The filament temperatures were obtained directly from the heating currents with the help of Langmuir's tables, 'Gen. Elect. Rev.,' vol. 30, p. 310 (1927).

cm./sec. for potassium, under the conditions of the experiments; (3) the asymmetric structure of the molecule, and its slow rate of temperature rotation, are favourable to the observation of intramolecular diffraction effects: assuming always the possibility that the Maxwell distribution of velocities in the primary beam, and the background of incoherent scattering, do not together mask the diffraction pattern.

Preliminary experiments were carried out in order to determine the correct temperature of the scattering oven O_2 to give single scattering of the alkali metal atoms in the *trans*-$C_2H_2I_2$ stream.

The primary beam was maintained steady, and the scattered intensity at a fixed angle was measured at a series of temperatures of the scattering oven.

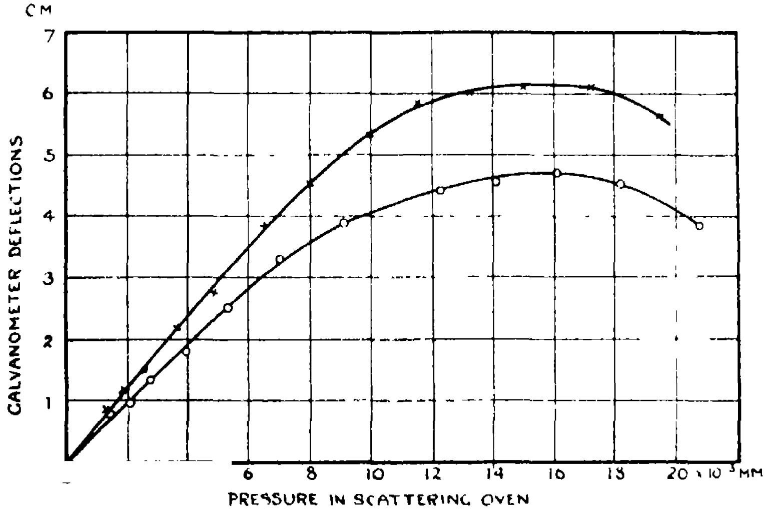


FIG. 3.—× = sodium, O = potassium.

In fig. 3, the scattered intensity, expressed as galvanometer deflections, is plotted against the pressure of $C_2H_2I_2$.* It can readily be shown that the departure of the curve from the initial exponential form, which extends from $p = 0$ to $p \simeq 14 \times 10^{-3}$ mm., indicates the onset of multiple scattering; the pressure in the scattering oven O_2 was held well below this point, namely, at $p = 7 \times 10^{-3}$ mm.

The distribution of scattering with angle was determined by traversing the detector system once in each direction. The scattered intensity was observed at each complete turn of the screwdriver. The measured intensity at a given position of the ring did not vary on the average by more than 2% for the forward and backward journey. Thereafter the oven O_2 was frozen out and the unavoidable background scattering was determined from point to point, and subtracted from the corresponding values of the total scattering. The background scattering, which we could show arose from stray atoms reflected from metal parts of the apparatus at room temperature, amounted on the average to 25% of the total scattering with potassium, and 15% with sodium. It was, however, always steady during a run to within 1% and hence does not affect the accuracy of the results.

---

* The vapour pressure curve of *trans*-$C_2H_2I_2$ has been determined by the effusion method (Broadway and Fraser, ' J. Chem. Soc.,' p. 429 (1933)).

Fig. 4 shows the angular distribution curves for the scattering of sodium and potassium beams, of the same original intensity, in *trans*-$C_2H_2I_2$ vapour.

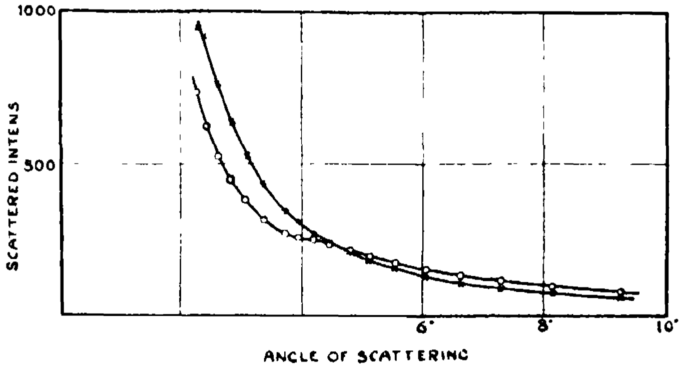


FIG. 4.—× -- sodium, T — 673° K.; O = potassium, T = 600° K.

In the figure, the scattering per unit solid angle, in arbitrary units, is plotted against the angle of scattering; each point represents the average of four observations.

The origin of the maximum at 4·25° on the potassium curve is obscure; it can, however, be stated definitely that it does not arise in the geometry of the apparatus. It might be interpreted as a second order diffraction maximum, but, if so, it is not clear why the corresponding sodium curve should show a monotonic fall. Knauer,* however, has recently reported somewhat analogous effects in the scattering of helium in helium, and hydrogen in hydrogen.

The most striking general feature of the curves is the fact that they cross between 4° and 5°; moreover, extrapolation of the curves to small angles suggests that they will cross again between 0° and 2°. A satisfying qualitative explanation of this general behaviour appeared to follow from the quantum theory of collisions of Massey and Mohr, p. 434, who generously gave us the results of their calculations considerably prior to publication. If one assumes as a first crude approximation that the alkali metal atoms and the $C_2H_2I_2$ molecule behave in collision as rigid spheres, and further that the radius of the sodium atom is less than that of the potassium, then the theory does, in fact, reproduce in rough outline the course of the experimental curves.

---

* 'Z. Physik,' vol. 80, p. 93 (1933).

It was decided therefore to test the theory as directly as possible, and the investigation described in the following paper, p. 634, was undertaken with this aim. The present scattering curves can best be discussed in conjunction with the additional experimental material presented there.

The present work on molecular scattering in gases is being financed by Imperial Chemical Industries, Limited ; thanks are due to Professor Lowry for giving it the ready hospitality of his laboratory. We are glad to record here our appreciation of Mr. S. A. McKay's skill in constructing apparatus to the rather exacting requirements demanded by the technique. One of us (L. F. B.) is indebted to the Department of Scientific and Industrial Research for a Senior Research Grant, during the tenure of which these experiments were carried out.

Summary.

A direct method of studying collision phenomena in gases, using crossed molecular beams, is described.

As a first application of the method, the angular distribution of the scattering of molecular beams of sodium and potassium in a stream of *trans*-di-iodo-ethylene vapour has been determined over a range of angle of 2°–10°. The results are in qualitative agreement with those predicted by the collision theory of Massey and Mohr.

Experiments on Molecular Scattering in Gases. II.—*The Collision of Sodium and Potassium Atoms with Mercury.*

By L. F. Broadway, Clare College, Cambridge.

(Communicated by T. M. Lowry, F R S.—Received March 28, 1933.)

An important parameter in the quantum theory of molecular scattering as developed by Massey and Mohr* is a quantity Q which corresponds roughly to a collision cross-section in classical theory. Of the quantities related to Q, that which is most directly accessible to observation is the decay constant of the intensity of a molecular beam traversing a scattering medium; the reciprocal of this constant corresponds to the mean free path of classical theory. Thus if I is the intensity of a beam which has passed through a distance d of a scattering medium, and I_0 is the original intensity, measured at the same distance from the source as is I, then

$$I = I_0 e^{-d/\lambda}, \tag{1}$$

where $1/\lambda$ is the decay constant; and we define Q by the relation

$$\lambda = \frac{1}{Q\nu_2\sqrt{1+\frac{m_1}{m_2}}} \tag{2}$$

where m_1, m_2 are the masses of the scattered and scattering molecules respectively, and ν_2 is the molecular concentration of the scattering medium.

The intensity I in equation (1) refers to those molecules which have suffered no deviation whatsoever in traversing the scattering medium. It is justifiable to assign this meaning to I, since quantum theory shows that if the interaction energy vanishes at infinity faster than $1/(\text{distance})^2$, the ratio I/I_0, and hence Q, is finite. Direct observation of the value of I/I_0 is, however, impossible, since it presumes an apparatus of infinite resolving power. Nevertheless, by using long narrow slits to define the beam, it is quite feasible to obtain angular apertures of only some 10^{-4} radians without serious loss of intensity†; and it is therefore pertinent to enquire how far the theory allows the deduction of the desired value of I/I_0 at zero angle from measurements made with beams of ribbon cross-section possessing small but finite apertures.

---

* Massey and Mohr, p. 434.

† *Cf.* Stern, 'Z. Physik,' vol. 39, p. 751 (1926).

Fig. 1, which is adapted from a diagram of Massey and Mohr (*loc. cit.*), illustrates the approximate form of the angular distribution curves to be expected on quantum theory from the collision of hard spheres, the sum of whose radii is $r_0 = r_1 + r_2$, and whose relative velocity is v. The constant k is given by $k = 2\pi Mv/h$, where $M = m_1m_2/m_1 + m_2$ is the "reduced mass." Curve I is the plot of the scattered intensity $I(\theta)$ per unit solid angle against the angle of scattering θ in relative co-ordinates. Curve II is derived from it by multiplying each ordinate by the appropriate value of $\sin\theta$, and represents therefore the scattered intensity *per unit angle* as a function of the angle of scattering; in other words, it corresponds directly to the lateral scattering of a beam of ribbon cross-section.

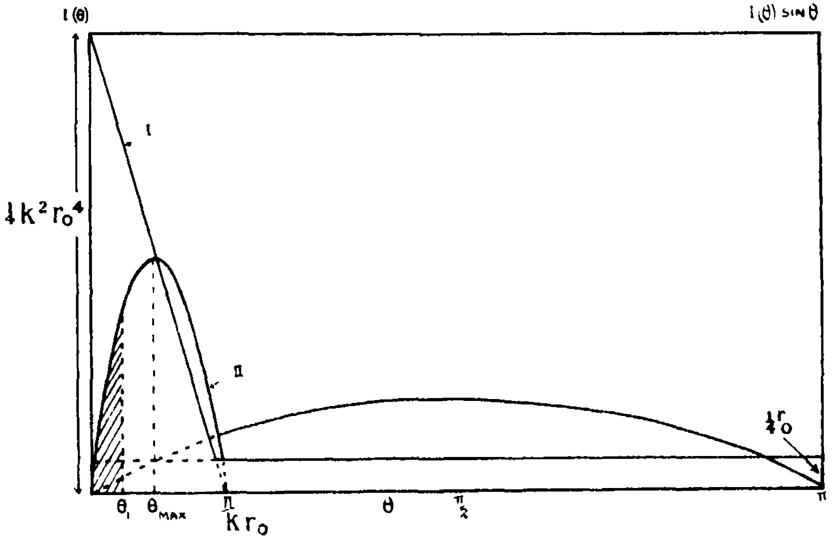


FIG. 1.

If now one attempts to measure I with a beam and detector system of finite aperture (say θ_1), one measures not only I, but also a proportion of the scattered molecules represented by the shaded area of curve II; that is, one measures a quantity $I'(\theta_1) = I + \int_0^{\theta_1} I(\theta) \sin\theta \, d\theta$. Thus a reliable evaluation of I demands the determination of $I'(\theta)$ *for a series of values* of the total aperture θ, ranging downwards from $\theta > \theta_{\max}$ to $\theta \leqslant \theta_{\max}$. The angle $\theta_{\max}$ depends on M, r_0, and v; a rough estimate of its value is therefore a necessary preliminary to the design of apparatus which shall be suitable for the determination of Q in any particular case.

The Decay of Alkali Metal Beams in Mercury Vapour.

The present experiments are concerned with measurements of the decay constants of beams of sodium and potassium in mercury vapour.

Mercury is a particularly suitable substance to use as scatterer in experiments designed primarily to test the applicability of a theory of the collision of hard spheres. Thus the vapour is monatomic, and the atoms can to a close approximation be regarded as spherical. The atoms are heavy compared with those of the alkali metal beam, which allows the angular distribution relative to the direction of the alkali metal beam to be compared directly with a theoretical distribution which is expressed in relative co-ordinates. The mercury atom is unlikely to possess a large attractive field towards the alkali metals, such as would invalidate the conclusions of the simple theory. Finally, the classical collision radius of the mercury atom is well known from viscosity measurements; hence its contribution to the measured interaction radius can be derived with some confidence from the theory.

In estimating $\theta_{\max}$, r_0 for the pair (K, Hg), which one could safely assume to be greater than r_0 (Na, Hg), was taken to be 5 A.; v was taken to be 5×10^4 cm./sec., giving $\theta_{\max} = 0 \cdot 43°$. Therefore a minimum total aperture of 0·2°, and a possible range of angular apertures of from 0·2° to 1°, were to be aimed at.

Apparatus.—It was evident that a very slight modification of the apparatus used in former experiments* would suffice to realize these requirements. The circular apertures defining the primary beam (*q.v.* I, fig. 1) were replaced by slits, 2 mm. high by 0·03 mm. wide. The canal forming the exit of the scattering oven (I, p. 629) was made 1 mm. in diameter and 1 mm. long, a form which allows direct use to be made of the angular distribution of the effusing atoms as calculated by Clausing† , its exit was situated 3 mm. below the centre of the primary beam. The positive electrode of the detector system (I, p. 627) now took the form of a vertical tungsten spiral, about 2 mm. diameter, masked by a slit, 2 mm. by 0·6 mm., placed some 15 mm. in front of it. The necessary variation of angular aperture could be effected with great ease in a single run by traversing the detector system as in previous experiments (I, p. 627).

A general view of the beam system is given in fig. 2.

---

* Fraser and Broadway, p. 626. Cited hereafter as I.

† 'Z. Physik,' vol. 66, p. 471 (1930).

Experimental Procedure.—I_0 was determined at a series of positions of the detector, with the mercury oven frozen out. Next $I'(\theta)$ was determined at the same series of positions, with a known pressure of mercury in the scattering oven. Finally, as a check on the constancy of the primary beam, the measurements of I_0 were repeated. In the best experiments with potassium, the values of I_0 repeated to within $\frac{1}{2}\%$; owing to the slow evaporation of the oxygen film from the filament (I, p. 630), the measurements for sodium were not so accurately repeatable.

Results.—In fig. 3 $I'(\theta)/I_0$ is plotted against the maximum horizontal deviation corresponding to a given aperture. Actually the detector slit, fig. 2, receives, in addition to all atoms scattered within this angle, an increasingly small percentage of atoms scattered at larger angles. Nevertheless, we may as a first approximation take $I'(0) = I + 2\int_0^\theta I(\theta)\sin\theta\, d\theta$; when differentiation of the curves of fig. 3 yields the corresponding angular distribution curves of $I(\theta)\sin\theta$ against θ, which are illustrated in fig. 4. They should be compared with the theoretical curve II of fig. 1.

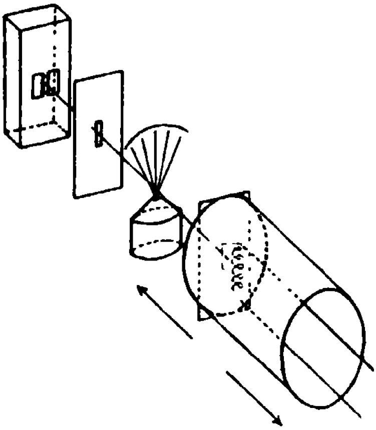


FIG. 2.

The curves of fig. 4 reproduce the general character of the theoretical curve so faithfully that we are justified in extrapolating them to zero angle. A criterion is thus at once provided for the extrapolation of the curves of fig. 3 to zero angle, so as to have a tangent parallel to the θ axis at $\theta = 0$. The value of $I'(\theta)/I_0$ for $\theta = 0$ so obtained is probably correct to 1%. This value is, however, simply the I/I_0 of equation (1), which is the starting-point in evaluating the quantum collision area Q.

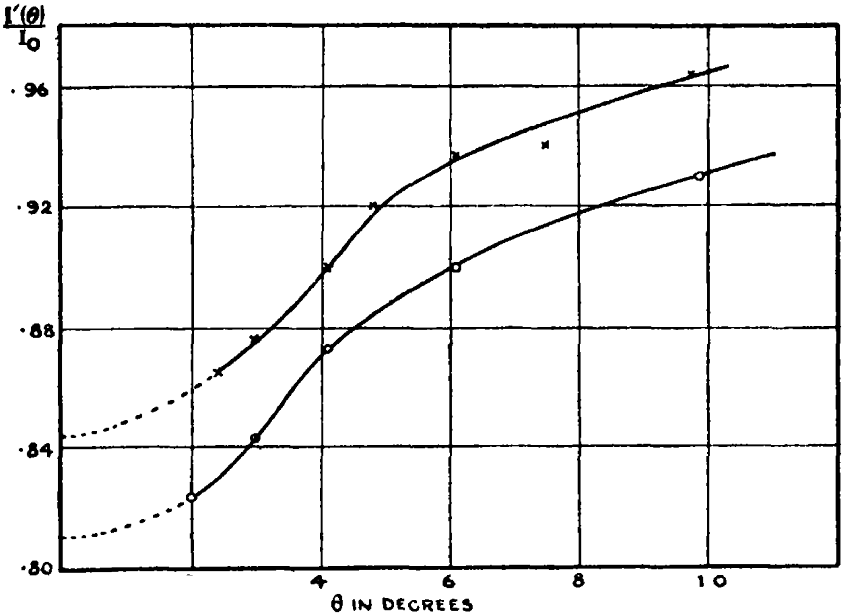


FIG. 3.

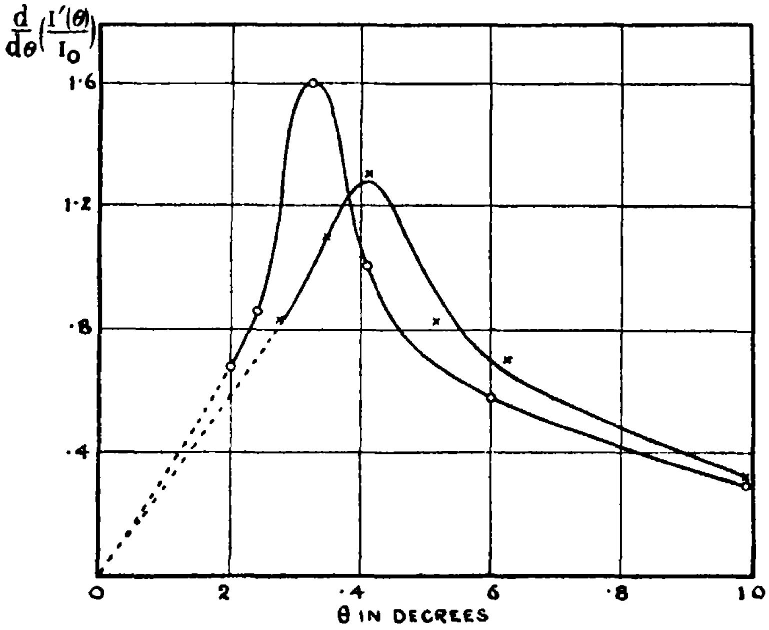


FIG. 4.

× = sodium, T = 616° K ; O = potassium, T = 543° K.

Evaluation of Q.—The molecular concentration in the stream of mercury atoms effusing from the orifice of the scattering oven, fig. 2, is not constant along the path of the primary beam; hence equations (1) and (2) require modification. It can readily be shown, from the calculations of Clausing (*loc. cit.*) that they are to be replaced by the equations

$$\mathrm{I} = \mathrm{I}_0 \exp. - \frac{\nu_0 \mathrm{R}^2}{2\kappa h} \cdot \int_0^{\pi/2} \mathrm{T} \cos \theta \, d\theta. \tag{3}$$

$$\kappa = \quad \mathrm{Q}\sqrt{1 + \frac{m_1}{m_2}} \tag{4}$$

where ν_0 is the concentration of mercury atoms in the oven; R is the radius of the orifice; h is the distance, assumed to be large compared with R, of a beam of infinitesimal cross-section above the exit of the oven; and T is a function of the radius R and length L of the exit canal, and of the direction of effusion θ. T has been tabulated by Clausing for 2R = L, which corresponds to the form of the exit canal used here; and the integral can be evaluated graphically from his data: $\int_0^{\pi/2} \mathrm{T} \cos \theta \, d\theta = 0 \cdot 641$ if 2R = L.

In an actual beam of small width and a height of a few millimetres, such as was used in the present experiments, it can be shown that equation (3) is applicable, with an error of considerably less than 1%, if h is taken to be the height of the centre of the beam above the exit. A practical limit is placed on the condition $h \gg \mathrm{R}$ by the necessity of obtaining a measurable decay of the primary beam in passing through a vapour stream of small cross-section which shall still maintain the essential condition of molecular effusion. The largest value of h allowed by these limitations was about 3 mm., corresponding to $h/\mathrm{R} = 6$; the error so introduced into the evaluation of Q was estimated from Clausing's calculated data to lie between 5 and 10%.

The values of Q for sodium and potassium in collision with mercury were found to be $3 \cdot 9 \times 10^{-14}$ cm.^{2} and $4 \cdot 7 \times 10^{-14}$ cm.^{2} respectively; the corresponding interaction radii are thus r(Na, Hg) = 11·1 A., r(K, Hg) = 12·2 A.

Conclusions.—These values are considerably larger than the classical collision radii r_0 estimated from the sum of the individual "viscosity" radii r_1 and r_2. Thus, for example, the values of r_1(K) and r_2(Hg) from estimations of electronic diameters and from viscosity measurements are 3·0 A. and 1·8 A. respectively, giving r_0(K, Hg) = 4·8 A., as compared with r(K, Hg) = 12·2 A. found above. This disparity in the values of the interaction radii determined by scattering and viscosity measurements respectively is in line with the predictions of quantum theory. Thus Massey and Mohr find that $r = \sqrt{2}r_0$ for zero attractive field; a weak attractive field would have the effect of increasing the ratio r/r_0 still further.

The results of the present experiments are qualitatively in harmony with the observations on the scattering of sodium and potassium in *trans*-di-iodoethylene vapour reported in I. Thus since $r(\mathrm{Na, Hg}) < r(\mathrm{K, Hg})$, it is fair to assume that $r_0(\mathrm{Na, C_2H_2I_2}) < r_0(\mathrm{K, C_2H_2I_2})$; moreover $k_{\mathrm{Na}} < k_{\mathrm{K}}$ under the conditions of the experiments; hence it is clear from the form of the approximate $\mathrm{I}(\theta):\theta$ curve of fig. 1 that the angular distribution curves for sodium and potassium scattered in *trans*-$C_2H_2I_2$ vapour should cross twice. The experimental curves are illustrated in fig. 4 of I; they do, in fact, cross once, in the sense predicted, within the observed range of angle; and extrapolation to smaller angles suggests that they will cross a second time.

Discussion.

Evaluation of I/I_0 at zero angle allows the determination of decay constants (or mean free paths) to be made independent of the geometry of the apparatus, and makes it possible to ascribe to them a definite meaning. The determination of I/I_0 requires a knowledge of the variation of $\mathrm{I}'(\theta)/\mathrm{I}_0$ with angle, and this in turn demands a restricted scattering volume in order that the angle of deviation may be rendered definite. Thus the method of measuring mean free paths, using the molecular rays technique, which has hitherto been used,* whereby a beam is scattered through a distance of several centimetres and the intensity diminution measured with a fixed aperture, is undoubtedly less well adapted to obtaining significant data than the method of crossed beams adopted here.

An extension of the measurements already made by this method is planned to include conditions where strong attractive fields may be expected to influence the collision process. A study of such cases cannot fail to give new information for theory to work on, and will probably lead to a better understanding from the experimental side of the primary processes of chemical reaction.

I am indebted to the Department of Scientific and Industrial Research for a Senior Research Grant, during the tenure of which these experiments were carried out; also to Professor Lowry for his sympathetic interest in the progress of the work.

---

* *Cf.* Knauer, 'Z. Physik,' vol. 80, p. 80 (1933).

Summary.

The decay of the intensity of molecular beams of sodium and potassium in a stream of mercury vapour has been measured as a function of the angular aperture of the beam system. It is shown that the experimental decay curves may be extrapolated unambiguously to zero angle, if use is made of the collision theory of Massey and Mohr. A determination of a decay constant (or mean free path) which is independent of the geometry of the apparatus is in this way rendered possible.

The Emission of Electrons from Tungsten and Molybdenum under the Action of Soft X-Rays from Copper.

By J. Bell, B.Sc.*

(Communicated by R. H. Fowler, F R S.—Received April 13, 1933).

1. *Introduction.*

The efficiency of many elements as emitters of soft X-rays has been investigated by Richardson and Robertson† and others, using a photoelectric method. The X-ray exciting voltages used ranged up to 6000 volts and several metals were used as photoelectric detectors. Some phenomena observed in large thermionic valves indicated a need for information about the photoelectric emission from tungsten and molybdenum in particular, under the action of X-rays from copper at rather higher exciting voltages than those previously studied, and for this investigation it was decided to follow the technique developed by Professor Richardson and Mr. Robertson to whom I am indebted for helpful discussion of the subject and guidance in the design of the apparatus.

The present results, therefore, form a continuation of the earlier work and extend it in several directions, principally the following: (1) the voltage range goes up to 20,000 volts; (2) tungsten and molybdenum were used as photoelectric detectors for X-rays from copper; and (3) an absorption screen, included in our apparatus, gave some indication of the quality of the radiation most effective in producing the emission. The effect of heat treatment of the photoelectric plates was also investigated and showed that the photoelectric emission is sensitive to the surface condition, but whether heat treatment produces an increase or a decrease in the emission depends on the previous treatment of the metal.

---

* Research Staff of the M.O. Valve Company, Ltd.

† 'Proc. Roy. Soc.,' A, vol. 115, p. 280 (1927); vol. 124, p. 188 (1929).

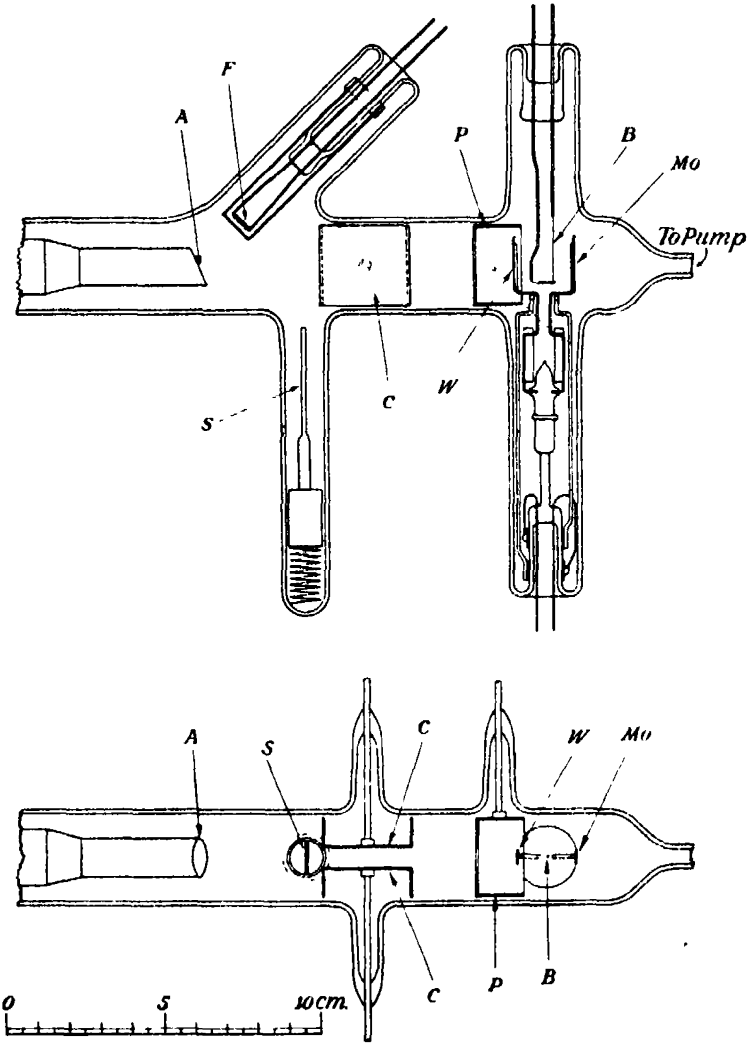


Fig. 1*a*.—Apparatus.

2. *Apparatus.*

The apparatus, which is shown diagrammatically in fig. 1*a* was in general design very similar to that of Richardson and Robertson (*loc. cit.*, 1927). A hard glass envelope with molybdenum seals was used thus eliminating all waxed joints. The X-ray tube filament (F) consisted of a narrow spiral of 4 ampere tungsten wire enclosed in a molybdenum focussing shield, and the anticathode (A), which was water-cooled, was made of well polished copper and sealed directly to the glass. The X-ray tube had a line focus, the line being in the plane of the slot between the condenser plates. The electron bombardment and the beam of X-rays selected were coplanar and were each at an angle of 11° on either side of the normal to the anticathode. The condenser plates (C), which were 3 cm. long by 2·9 cm. broad separated by a gap of 0·5 cm., were made of nickel and the ends were bent round to fit the glass tube closely, thus ensuring that the only passage down the tube was through the gap between the plates. Between anticathode and condenser plates was a side tube containing an aluminium window (S) 0·02 mm. thick mounted on an iron base, so that by means of a magnet it could be raised into the X-ray beam to study absorption effects. The anode for the photoelectric detector (P) consisted of a nickel cylinder, the front of which was closed except for a slot slightly wider than the gap between the condenser plates. The photoelectric plates themselves, one of tungsten and one of molybdenum were 2 cm. long and 0·5 cm. wide, they were mounted on a short length of glass tubing closed at one end, which was pivoted on a glass rod. This system, then, could be rotated by the action of magnets on soft iron pieces fixed to the photoelectric plate supports, and either plate placed and located in a definite position inside the anode. In this position the photoelectric plate was at a distance of 10 cm. from the anti-

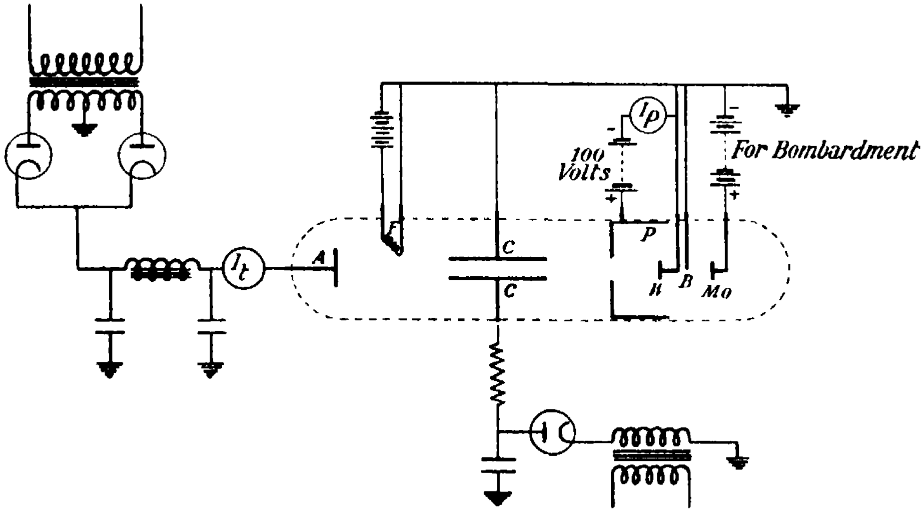


FIG. 1*b*.—Electrical connections.

cathode. With one plate inside the anode, the other was in position for heating by electronic bombardment from a small filament (B). The bombardment was confined to the inside of the plate, hence the face which was later to be exposed to the X-rays did not become contaminated with tungsten evaporated from the filament.

The design of the tube was such that the whole focal line was visible from every point on the photoelectric plate, this facilitated calculation of the solid angle subtended by the plate at the anticathode. The tube was connected to a 3-stage mercury diffusion pump through a liquid air trap and mercury cut-off, and before taking any readings of photoelectric emission was baked out in an electric oven at 500° C. for 1 hour.

Fig. 1*b* shows the electrical connections. The high tension supply for the X-ray tube was obtained from a transformer whose output was rectified and smoothed. The transformer primary was fed from a 300-cycle alternator and continuous voltage variation obtained by potentiometer control of the alternator excitation. The anticathode D.C. voltage was measured by a milliampere meter in series with an accurately known high resistance. The condenser plate potential, which was obtained from a single phase rectified and smoothed supply, was measured with an electrostatic voltmeter. Throughout the experiments, the anode of the photoelectric detector was maintained at about +100 volts with respect to the photoelectric plate whose emission was being studied. In general the X-ray tube current was adjusted to about 10 milliamperes, this value was chosen as it gave a photoelectric current from the plate of magnitude suitable for measuring with a sensitive galvanometer. The lead from photoelectric plate to galvanometer was screened electrically, and was also screened from X-rays in the vicinity of the tube by means of lead sheet.

3. *Experimental Results.*

The ultimate object of the experiments was to investigate the photoelectric emission from heat treated tungsten and molybdenum at X-ray exciting voltages in the range 1 to 20 kv., but it was also desired to study the variation in emission after the photoelectric plates had been heated at various temperatures.

Before assembly in the apparatus the tungsten and molybdenum photoelectric plates were not heat treated, the only cleaning being chemical. It is important to keep this in mind as it has a bearing on the emission in the early stages of heating. It was necessary first to determine the voltage required on the condenser plates to filter out all charged particles from the X-ray beam. Fig. 2 shows the variation of current from the photoelectric plate with condenser plate potential difference, for X-ray exciting voltage = 12 kv. It will be seen that below about 1000 volts the photoelectric plate current is dependent on the particular value of condenser plate potential, which indicates that charged particles are affecting the photoelectric response. Beyond 1000 volts, however, the photoelectric plate current is steady, implying that no charged particles are reaching the plate and that the steady current is due solely to X-rays. In the subsequent experiments the condenser plate voltage was fixed at one-sixth of the X-ray tube voltage and this was found to give an ample margin of safety.

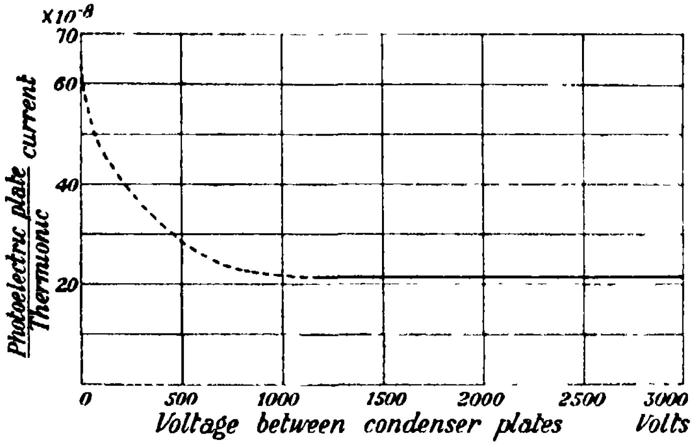


FIG. 2.—Voltage between condenser plates required to remove charged particles from the X-ray beam. Exciting voltage 12 kv.

Curves of $\frac{\text{Photoelectric current } (I_p)}{\text{X-ray exciting current } (I_t)}$ against X-ray exciting voltage, for molybdenum and tungsten, before any heating other than the 500° C. bake, are shown dotted in fig. 3. It was confirmed that I_p/I_t was independent of I_t over the range of I_t used. These curves were not taken beyond 13 kv. for tungsten and 14 kv. for molybdenum, but throughout this range of voltage the photoelectric emission increases gradually as the X-ray exciting voltage is increased. The hump on each curve between 5 and 8 kv. was found subsequently to be caused in part by gas contamination arising from the X-ray tube end of the apparatus, this is discussed in more detail below. The heat treatment of the photoelectric plates was carried out in stages, the photoelectric current being measured immediately after the metal had been heated at each temperature. It was found that 45 minutes was sufficient heating time, that is, further heating at the same temperature produced no change in the photoelectric emission. In the first experiments it was found that after heating the plates the photoelectric current was rather unsteady, tending to decrease with time especially in the voltage range 5 to 8 kv. This effect was traced partly to the X-ray tube end of the apparatus and seemed to be caused by adsorption of gas on the anticathode. However, it was not determined whether the increased photoelectric emission was caused by an enhancement of the soft radiation or by slight gas contamination of the photoelectric plate when gas was bombarded off the anticathode. The evidence pointed to a combination of the two effects. In order to minimize the effect as far as possible, the X-ray tube was operated during the time the photoelectric plates were bombarded as this helped to prevent adsorption of gas on the anticathode. The effect of heat treatment of the photoelectric plates at this stage can be seen from the full line curves in fig. 3, and was to cause an increase in the photoelectric emission from both metals. The full line curves in fig. 3 were taken after each plate had been heated at about 1000° C. (The design of the apparatus did not permit of accurate measurement of the temperatures and the values given are estimated.) Further heating at higher temperatures

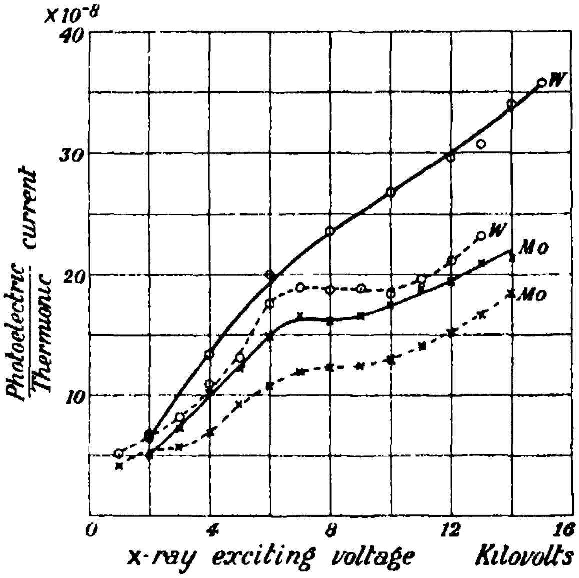


FIG. 3.—Dotted curves taken after baking at 500° C ; full line curves taken after heating plates to about 1000° C.

produced only slight change in the emission, the highest temperatures reached were about 1650° C. for molybdenum and 1800° C. for tungsten.

The final curves obtained after heating the photoelectric plates at high temperatures are shown in fig. 4; these curves represent the true X-ray photoelectric effect for well degassed tungsten and molybdenum. The complete effect of heat treatment and degassing can be seen from fig. 5, this is for a tungsten photoelectric plate at X-ray exciting voltage 10 kv. It will be seen that the photoelectric current can vary by a factor of about 2 depending on the surface condition of the metal.

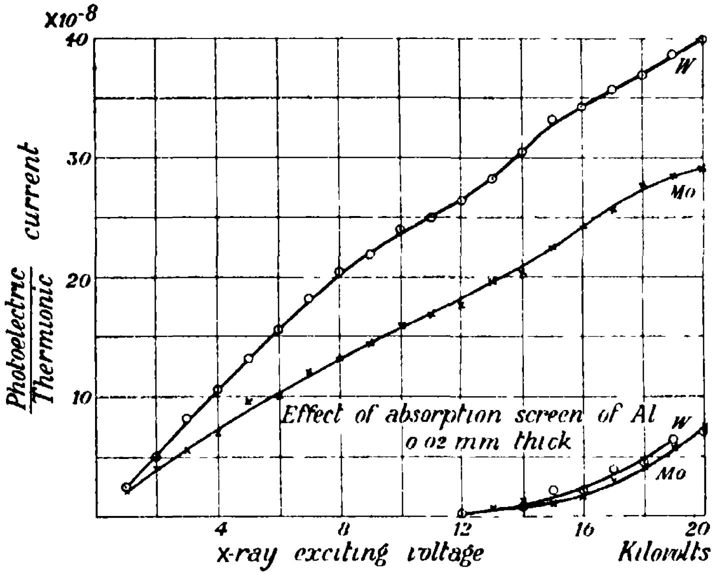


FIG. 4.—Final curves after plates had been heated to very high temperatures.

The increase of photoelectric current on heating the photoelectric plates seen in fig. 3 is contrary to the result recorded by Nakaya† who found a decrease of the emission after heat treatment. It was thought that this difference might be accounted for by some effect of the chemical cleaning in the present case, and that the increase of photoelectric current after heat treatment corresponded with the removal of a chemical residue. In order to confirm this point the tube was opened to air for about 20 hours after thorough heat treatment of both plates, and, after a further bake at 500° C., it was found that the photoelectric emission was higher than at any time during the previous experiments. In addition, heating now produced a decrease in the emission, in agreement with Nakaya's results (see fig. 5). Both photoelectric plates were again heated at high temperatures and the curves extended to 20 kv. These are the curves of fig. 4. The photoelectric emission from both metals is still increasing at 20 kv. and there is no indication of any inherent change of slope at immediately higher voltages. Assuming a uniform distribution of the radiation, we may obtain the photoelectric currents which would have been obtained if the total radiation generated at the anticathode had been effective at the photoelectric plate by multiplying the observed values by 628.

---

† 'Proc. Roy. Soc.,' A, vol. 124. p. 616 (1929).

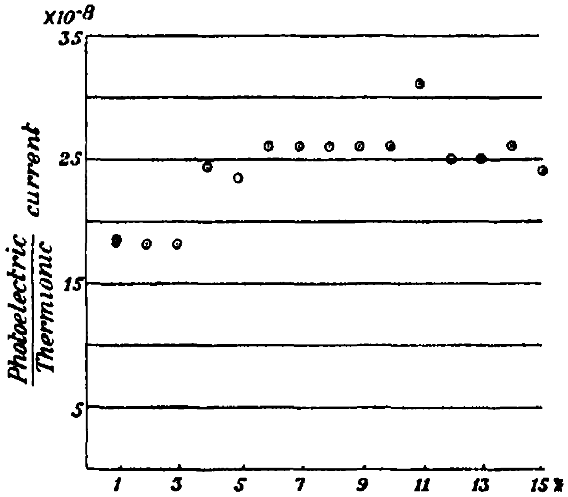


FIG. 5.—Tungsten photoelectric plate. X-ray exciting voltage 10 kv. Complete effect of heat treatment and degassing. *Remarks on treatment:—(1) Three days after bake at 500° C.; (2) seven days after bake at 500° C.; (3) eight days after bake at 500° C.; (4) after very slight heat treatment; (5) after bombardment at 10 watts; (6) after bombardment at 28 watts; (7) after bombardment at 38 watts; (8) after bombardment at 45 watts; (9) after bombardment at 67 watts; (10) after bombardment at 80 watts; (11) tube opened to air + bake at 500° C. + two days; (12) after bombardment at 10 watts; (13) after bombardment at 30 watts; (14) after bombardment at 60 watts; (15) after bombardment at 80 watts.

An interesting point arises from the curves taken at various stages of heat treatment; this is that heating does not affect relative values of photoelectric emission from tungsten and molybdenum. In the completely unbombarded condition the emission from molybdenum was about 70% that from tungsten at the same X-ray exciting voltage, also this ratio was maintained after both metals had been heated at the same temperature. Further, after the tube had been opened to air, producing a large increase in the emission, the current from molybdenum was still some 70% that from tungsten. This result would tend to show that even when the surface of the photoelectric plate is contaminated, either by gas or chemical residue, the photoelectrons come from the metal itself and not from the surface film. The question whether the addition of a surface film to an already clean metal surface will enhance or suppress the emission should depend on the nature of the film and this seems to be supported by experimental evidence. (1) Removal of chemical residue + gas film increased the emission, (2) gas contamination of the already clean surface increased the emission, (3) removal of this gas contamination reduced the the emission, fig. 5.

Experiments with an absorption screen of 0·02 mm. aluminium in the X-ray beam, fig. 4, gave some idea of the quality of the radiation which is most effective in producing photoelectrons. Below an X-ray exciting voltage of 10 kv. no photoelectric current could be detected with the galvanometer used (1 mm. $= 2{\cdot}83 \times 10^{-10}$ amperes). The effective absorption of the X-rays is thus very heavy implying that the photoelectric emission is due amost entirely to the very soft components of the radiation. This is to be expected owing to the greater penetrating power of the harder rays which produce photoelectrons deeper in the metal. At 20 kv. the absorption from the photoelectric point of view is about 80% and 20 kv. radiation with 0·02 mm. absorption screen is equivalent to about 4 kv. radiation without screen, hence for experiments of this kind it is not practicable to use an X-ray tube with a window. The above explains why the incidence of characteristic CuK radiation at about 8800 volts is not detectable on the photoelectric curves. To obtain further information about the precise radiation which is most effective from the photoelectric view point, additional thin absorption screens, which unfortunately were not included in the present apparatus, would have to be used. It is interesting to note that in an earlier experiment where several absorption screens of different thicknesses were used (the thinnest was 0·038 mm.) we found that the radiation, in the range 10 to 16 kv., appeared surprisingly homogeneous from the photoelectric point of view. This, of course, is neglecting the radiation absorbed by the 0·038 mm. window which was the window of the tube itself. A value for the apparent quality of the radiation could be deduced from knowledge of the thickness of absorber required to reduce the photoelectric intensity to half value, which was obtained from plots of the logarithm of the intensity against added thickness of absorber. This apparent quality was found to vary only slightly with exciting voltage between 10 and

16 kv. The radiation for 10 kv. exciting voltage corresponded to about 4·5 kv. quality and for 16 kv. exciting voltage to about 6·3 kv. quality.

Bandopadhyaya* concluded that the softer radiations in the beam of soft X-rays were most effective in producing photoelectrons and this is supported by our results with the absorption screen. Bandopadhyaya also studied the effect of degassing on many metals used as photoelectric detectors. He found that in general the photoelectric current was much higher with the photoelectric plate entirely undegassed but that degassing affected relative values. As mentioned previously, we found that heat treatment did not affect relative values for tungsten and molybdenum. It therefore appears that our result may be a coincidence and that it does not hold in general for other metals, but at the same time great care must be taken when comparing figures for degassed and undegassed metals to ensure that the undegassed states of the different metals are similar.

As far as can be ascertained, the qualitative agreement between our results and those of Richardson and Robertson is good. Taking figures from Richardson and Robertson's fig. 2 (*loc. cit.*, 1929), the rate of increase of photoelectric emission with X-ray exciting voltage can be found. For a copper anticathode and nickel photoelectric detector they find the emission at 6 kv. to be 2·2 times that at 2 kv. We find for a copper anticathode and molybdenum detector an increase of 2·6 times in the same range and for tungsten an increase of 3·1 times.

In conclusion, the author desires to tender his acknowledgment to The General Electric Company and the Marconi Company on whose behalf much of the work was done which has led to this publication.

He also wishes to thank Professor O. W. Richardson, F.R.S., and Mr. Robertson for a discussion of the subject before these experiments were commenced, and Professor R. H. Fowler, F.R.S., for his helpful interest in the work.

4. *Summary.*

The photoelectric emission from molybdenum and tungsten under the action of soft X-rays from copper in the voltage range 1 to 20 kv. is investigated.

The final curves for the clean metals after heating at high temperatures are given; the photoelectric emission increases regularly with increase of X-ray exciting voltage throughout the range studied. The effect of heat treatment is also dealt with; heating may produce either an increase or a decrease in the

---

* 'Proc. Roy. Soc.,' A, vol. 120, p. 46 (1928).

photoelectric emission depending on the previous treatment of the metal, but relative values for molybdenum and tungsten are not much affected by heat treatment. The emission from molybdenum is about 70% that from tungsten.

Experiments with an absorption screen of aluminium, 0·02 mm. thick, show that the photoelectric emission is produced mainly by the very soft components of the radiation.

The Flow Past Circular Cylinders at Low Speeds.

By A. Thom, D.Sc., Ph.D., Carnegie Teaching Fellow, University of Glasgow.

(Communicated by G. I. Taylor, F R S —Received April 24, 1933.)

[Plates 10–13.]

Introduction.

This paper deals chiefly with calculations and experiments on the flow past circular cylinders, but the arithmetical methods of solution of the equations of steady viscous flow proposed and used in Section I, are applicable to other equations and may be of interest.

Section I.—*Arithmetical Methods of Solving the Equations of Viscous Steady Flow.*

The equations to be solved are

$$\left.\begin{aligned} \nu\,\nabla^2\zeta &= \frac{\partial\psi}{\partial x}\frac{\partial\zeta}{\partial y} - \frac{\partial\psi}{\partial y}\frac{\partial\zeta}{\partial x} \\ \nabla^2\psi &= 2\zeta \end{aligned}\right\} . \tag{1}$$

The method of solution used* is one of repeated interpolation in a field of initially assumed values. As the formulæ of interpolation used involve the equations to be solved the values tend to settle to a solution provided certain precautions are taken. In practice the field is divided into squares and initially assumed values of ψ and ζ are written at each corner. In general the values at the centre of a square of side $2n$ are given by

$$\zeta_O = \zeta_M - \tfrac{1}{2}n^2\nabla^2\psi, \quad \psi_C = \psi_M - \tfrac{1}{2}n^2\nabla^2\psi, \tag{2}$$

---

* Thom and Orr, 'Proc. Roy. Soc.,' A, vol. 131, p. 30 (1931); Thom, 'Aer. Res. Ctee.,' R. & M., No. 1194 (1929).

where

$$\psi_M = (A + B + C + D) \div 4$$
$$\zeta_M = (a + b + c + d) \div 4,$$

A, B, C and D are the corner values of ψ and a, b, c and d are the corresponding values of ζ.

These expressions, which are easily obtained by applications of Taylor's Theorem, are correct up to and including ∂^3 terms. In the particular case under consideration, namely, the solution of equations (1) above we obtain :—

$$\left.\begin{aligned} \zeta_C &= \zeta_M - \tfrac{1}{2}\frac{n^2}{\nu}\left(\frac{\partial\psi}{\partial x}\frac{\partial\zeta}{\partial y} - \frac{\partial\psi}{\partial y}\frac{\partial\zeta}{\partial x}\right) \\ \psi_C &= \psi_M - n^2\zeta_C \end{aligned}\right\} . \qquad (3)$$

Writing approximate values for $\partial\psi/\partial x$, $\partial\zeta/\partial y$, etc., the final working form of these become :—

$$\left.\begin{aligned} \zeta_C &= \zeta_M - \frac{1}{16\nu}\{(a - c)(B - D) + (b - d)(C - A)\} \\ \psi_C &= \psi_M - n^2\zeta_C \end{aligned}\right\} . \qquad (4)$$

Having used these to find the values at the centres of all squares they are used again to find new values at the original corners. The process is repeated over and over again till the values all over the field have settled. At solid boundaries the values of ζ on the surface are obtained from the approximate expression :—

$$\zeta = (\psi_G - \psi_S) \div m, \qquad (5)$$

where ψ_S = value of ψ on the surface, and ψ_G = value of ψ at a point G distant m from the surface. This expression (or an equivalent) has to be used each time the squares are gone over.

If the squares used are too large the process may not be convergent, and even if convergent distorted results may be obtained.

As an example of the process consider the flow past a circular cylinder at Reynolds' Number $(Vd/\nu) = 10$. Part of the solution is shown in fig. 1. The particulars are : undisturbed velocity = 6·25, cylinder diameter = 10, coefficient of kinematic viscosity ν = 6·25. It will be seen that throughout the greater part of the field the ψ and ζ values are repeating to a reasonable accuracy, indicating that an approximate solution has been obtained. The streamlines thus obtained are shown in fig. 2. When the approximation is carried further there is an indication of negative values of ψ and ζ just behind the cylinder showing the presence of a small weak eddy. Further work on this field has, however, not been carried out as a different method of procedure presented itself.

It will be realized that considerable difficulty arises at a curved boundary because it is impossible to arrange matters so that the boundary passes through

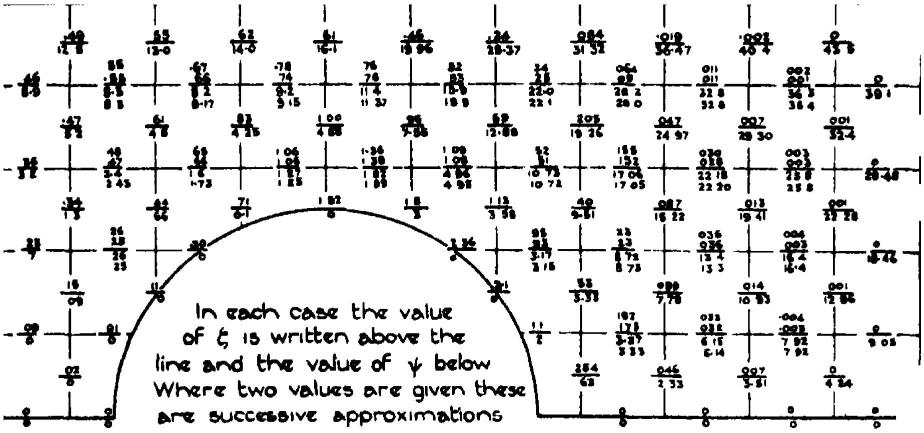

In each case the value of ζ is written above the line and the value of ψ below Where two values are given these are successive approximations

FIG. 1.

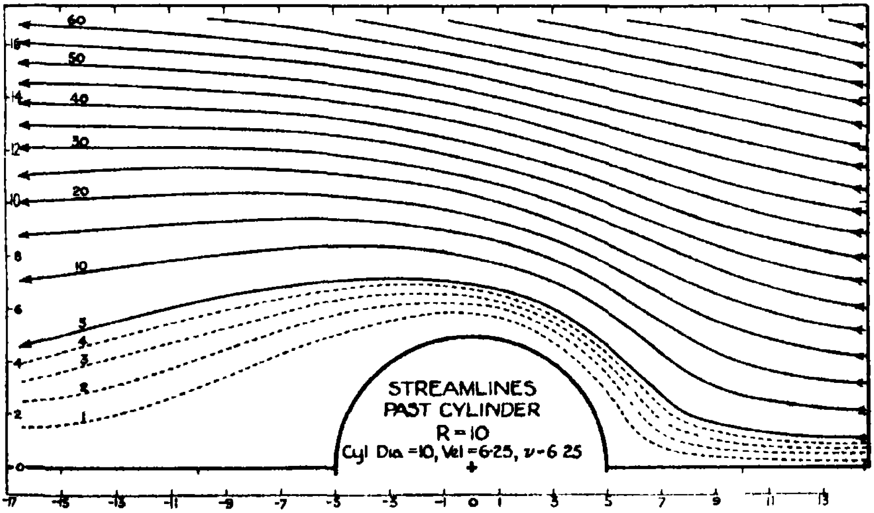


FIG. 2.

the corners of all the squares it cuts. Accordingly some method of interpolation has to be used to obtain the values of ψ and ζ for the successive approximations on the corners of those boundary squares whose outer corners fail to fall exactly on the boundary. With fields of this type (and naturally they are in a majority) there is also more difficulty in calculating the ζ values on the solid boundaries.

These difficulties are both overcome by the following method.

Graphically or otherwise solve the equation $\nabla^2\psi = 0$ for the given boundaries and so obtain the network of streamlines and equipotential lines dividing the field into " squares " of any desired size. These " squares " can now be used to obtain numerical solutions for the steady flow of a viscous fluid (and a variety of other problems also) in a similar manner to that already developed for the true squares of the xy network in the z plane.

There is now no trouble at the boundaries (apart from those inherent in the particular problem under solution) since the sides of the squares of the new grid lie everywhere on the boundaries. The proof of the validity of the process is as follows :—

Let the grid co-ordinates be ξ and η, *i.e.*, the lines forming the network have the equations in the xy field ξ = constant, and η = constant.

Put

$$z = x + iy$$

$$t = \xi + i\eta,$$

and write

$$\nabla_t{}^2\zeta \equiv \frac{\partial^2\zeta}{\partial\xi^2} + \frac{\partial^2\zeta}{\partial\eta^2},$$

so that the subscript after ∇ indicates the co-ordinates with respect to which the differentiations are carried out.

Let the transformation required to produce the grid be

$$t = f(z).$$

The equations (1) may now be transformed from the z to the t field when they become

$$\nu\nabla_t{}^2\zeta = \frac{\partial\psi}{\partial\xi}\frac{\partial\zeta}{\partial\eta} - \frac{\partial\psi}{\partial\eta}\frac{\partial\zeta}{\partial\xi}$$

$$q^2\nabla_t{}^2\psi = 2\zeta$$

These equations only differ from those for the z field by the q^2 in the second member, and hence the method of treatment already described applies to their solution. q is the velocity of transformation.

The working formulæ are

$$\left.\begin{aligned}\zeta_{\rm O} &= \zeta_{\rm M} - \frac{1}{16\nu}\{(a-c)({\rm B}-{\rm D}) + (b-d)({\rm C}-{\rm A})\}\\ \psi_{\rm C} &= \psi_{\rm M} - n^2\zeta_{\rm C}/q^2\end{aligned}\right\}, \tag{7}$$

where the subscript M indicates the mean of the corner values, small letters are corner values of ζ and capitals corner values of ψ. The side of the square in the t plane is $2n$. The true length of the side of the square in the z plane is $2n/q$ so that actually there is no difference whatever between these formulæ and those previously given. In working, however, it is inconvenient to use the original z plane with the grid drawn on. Instead the sheets are prepared for the t plane so that ordinary squared paper can be used (*e.g.*, see fig. 3).

When new corner values have been obtained, the values of ζ on the solid boundaries are revised from the formulæ

$$\zeta = (\psi_{\rm G} - \psi_{\rm S})\, q^2/m^2, \tag{8}$$

where $\psi_{\rm G}$ is the value of the stream function at a point G, distant m from the boundary in the t plane, that is at a distance m/q in the z plane, and $\psi_{\rm S}$ is the value on the boundary. In (8) the value of q to be used is the mean of the surface value and the value at G. To test that G is sufficiently near the surface for the process to be valid, apply (8) to two points on the same normal, one being twice as far from the surface as the other. If there is a serious discrepancy, then the distance even of the inner is presumably too large. For the first few rounds, however, until the values have begun to settle, it is better not to use too small a square as the work is thereby rendered very laborious, and in any case, the discrepancies may be due to the field not being settled.

The example chosen to illustrate the method is that of the flow past a circular cylinder in an infinite field at R = 20. For this, two transformations were available, namely, $t = \log z$ and $t = z + 1/z$. The first, giving as it does a system of radiating lines and concentric circles, has the advantage of having smaller squares near the cylinder and larger ones further out. The second was, however, chosen because of its similarity to the actual flow.

The " grid " having been plotted to a large scale and the values of q obtained the calculation of the viscous flow was commenced using the following values :—

d = cylinder diameter = 2

V = undisturbed velocity = 1

$\nu = 0{\cdot}1$

giving

$$R = Vd/\nu = 20.$$

To start the solution values of the vorticity ζ, and stream-function ψ were roughly estimated from the approximate solution for $R = 10$ already obtained. These having been written on the corners of squares in the t field, the work commenced. Formulæ (7) and (8) were applied until the field had settled.* In choosing the initial values no attempt was made to assume an eddy condition by writing negative values behind the cylinder. All the initially assumed values were positive. The negative values developed during the work, indicating a pair of symmetrical eddies in this region.

A portion of the work is shown in fig. 3. The sheet shown includes the eddy region. If a sample calculation is made it will be seen that no addition to the standard formulæ is required to deal with the eddy, in fact here the work is lightened as the second term in the formulæ (7) is so small in the eddy region as to be almost negligible, and $\nabla^4\psi = 0$ almost applies here. Fig. 3 also indicates how the size of square used varied, being generally larger the greater the distance from the cylinder. The criterion for the size of square is to try smaller squares and if serious differences are obtained, all the squares in the neighbourhood must be reduced in size.

It was found necessary to use very small squares in the region adjacent to the front generator—possibly because this is a singular point in the grid. The method of treating this region was substantially the same as that indicated elsewhere for the sharp intruding corner in the problem of the sudden expansion in a pipe.†

The whole process is long and no claim is made that the values obtained are perfect. The use of smaller squares everywhere would probably cause slight changes. It is, however, considered that the solution is better than that obtained previously for $R = 10$. The values of ψ and ζ obtained are shown plotted as contours in the t field in fig. 4. By means of the " grid " and some cross-plotting it is easy to obtain the actual streamlines. These are shown in fig. 5. The eddy is clearly shown in both figs. 4 and 5 It will be seen to correspond closely with the actual photographs of the eddy shown in Section 3. In size, it is nearest to the photograph at $R = 29$ (see fig. 12 (*d*), Plate 13).

---

* With regard to the number of repetitions required at a point, it is impossible to give a definite figure. It would depend on the boundaries, size of square, accuracy of the assumed solution, the accuracy of the required solution, and the position of the point in the field. Usually the number lies between 5 and 50.

† 'Aero. Res. Ctee.,' R. & M., No. 1475 (1932).

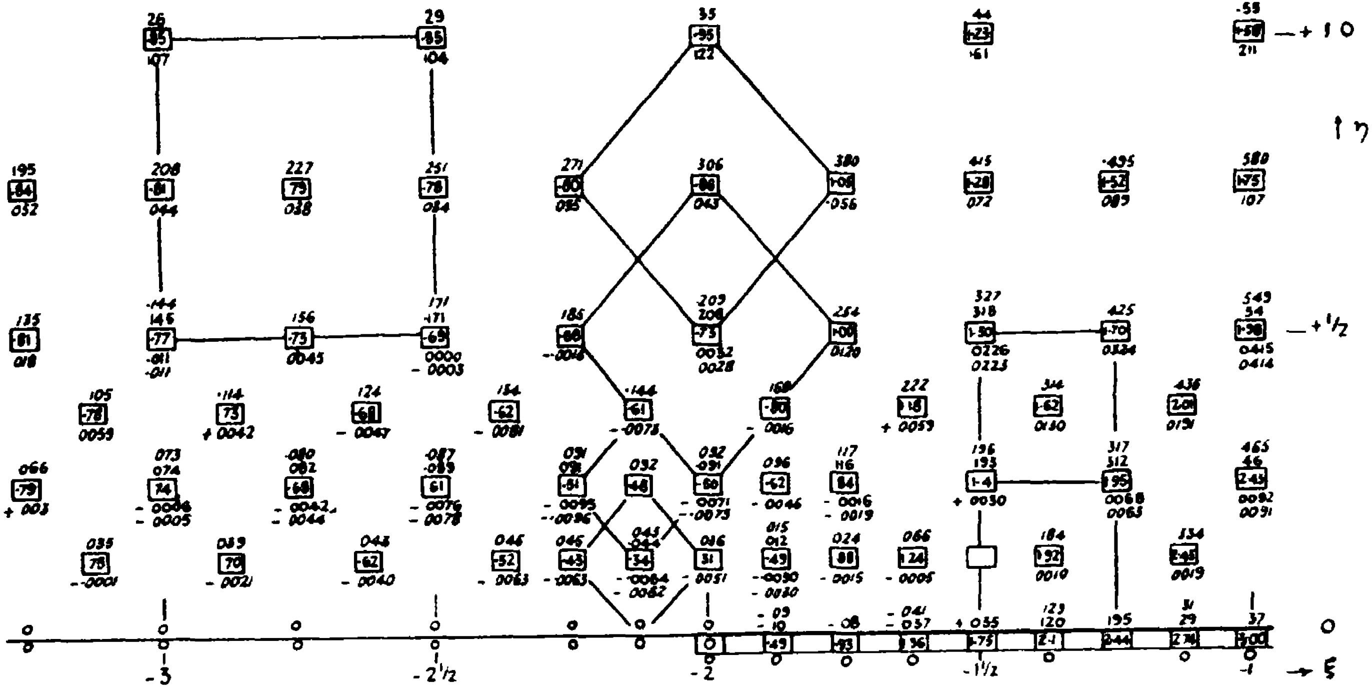


FIG. 3.—A portion of the working including the eddy region Figures shown thus 0·082 were first taken from a previous sheet; from these the sloping figures were calculated. Values of q^2 are enclosed in rectangles. Values of ζ are written above the rectangle and values of ψ below.

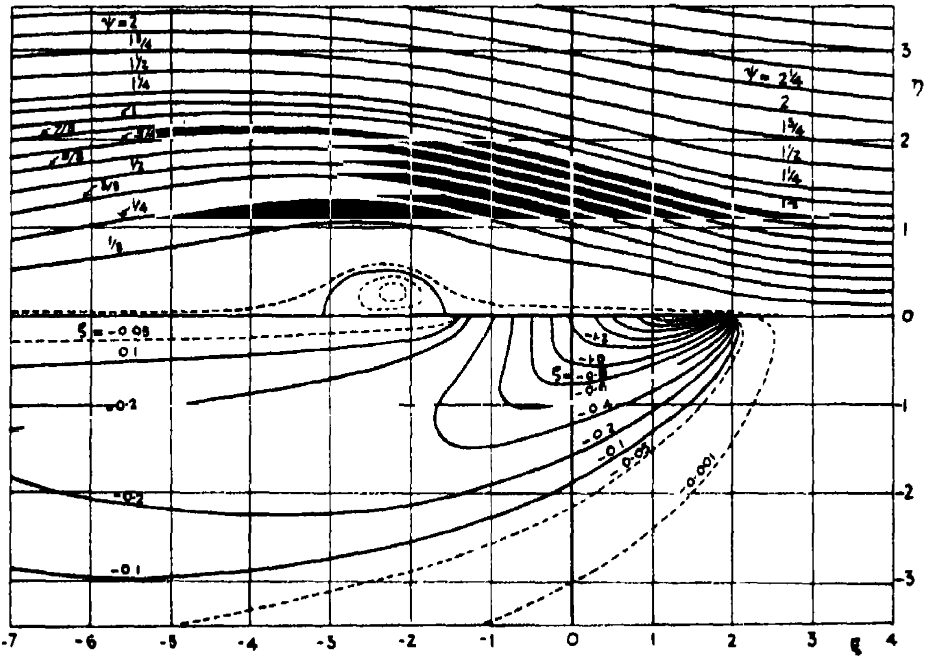


FIG. 4.—Streamfunction and vorticity contours in the t plane.

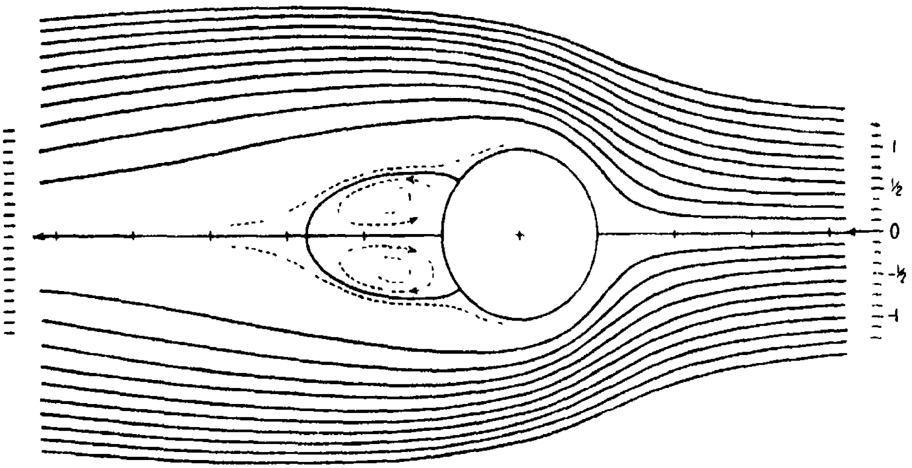


FIG. 5.—Calculated streamlines past a circular cylinder at $Vd/\nu = 20$. The spacing of the streamlines in the undisturbed flow is shown by the short lines at each end. The dotted streamlines are at $+ 1/256$, $- 1/256$, and $- 1/128$.

Fig. 5 can also be compared with the photograph obtained at $R = 23$ (see fig. 9, Plate 10). A further comparison with experiment is shown in fig. 13 where the maximum separation of the streamlines which were originally at a distance apart equal to the cylinder diameter is plotted on Reynolds' number. These were obtained from a number of photographs similar to fig. 9, Plate 10. It is seen that the value obtained from the theoretical streamlines is in good agreement. The distance behind the cylinder at which this maximum occurs is also shown in fig. 13 and the agreement is here again satisfactory.

To determine the pressure at any point A in the field it is necessary to integrate along a line from some point B where the pressure is already known to A.* In the present example B is conveniently a point at such a distance from the cylinder that the flow there is undisturbed by the cylinder. In deducing the necessary expressions for use in the t plane the factor q cancels out. For example Bernoulli's equation for use along a line parallel to the η axis is

$$p_A + \tfrac{1}{2}\rho {q_A}^2 = p_B + \tfrac{1}{2}\rho {q_B}^2 - 2\rho\nu \int_A^B \frac{\partial\zeta}{\partial\xi}\, d\eta + 2\rho \int_A^B \frac{\partial\psi}{\partial\eta}\, d\eta$$

The pressure on the front generator was obtained by integrating along $\eta = 0$, and also by integrating along $\xi = 2$ in the t field. The results are $1\cdot31$ ($\tfrac{1}{2}\rho V^2$) and $1\cdot35$ ($\tfrac{1}{2}\rho V^2$) respectively. The mean value $1\cdot33$ is in good agreement with the value ($1\cdot335$) obtained elsewhere theoretically by a totally different process.†

Pressures at other points on the cylinder surface were obtained by integrating along lines parallel to the η axis in the t field. The results are as follows :—

ξ.	ζ_0.	θ.	$(p - p_0)/\tfrac{1}{2}\rho V^2$.
2·0	0	0	+1·33
1¾	1·75	29·0	+0·67
1½	2·1	40·3	+0·12
1	2·0	60·0	−0·52
0	1·18	90·0	−1·16
−1½	0·03	139·7	−0·87
−2	0	180·0	−0·73

These values will be found plotted in fig. 6 among the experimental curves. The pressure drag is easily found. Expressed as a coefficient it is $K_D = 0\cdot624$.

---

* See R. & M., No. 1194 (1929).

† 'Aero Res. Ctee.,' R. & M., No. 1389, Table 5 (1931).

It has been shown elsewhere that the tangential force at a point is $2\mu\zeta_0$,* where ζ_0 is the value of the vorticity on the surface. The values of ζ_0 are given in the above table. Integrating the drag component of the tangential force and expressing the result as a coefficient gives $K_D = 0{\cdot}433$. Thus the total drag coefficient is $1{\cdot}057$ at $R = 20$, a result lying between the experimental values of Relf and those of Wieselberger. These results are shown on fig. 8 plotted along with experimentally determined values.

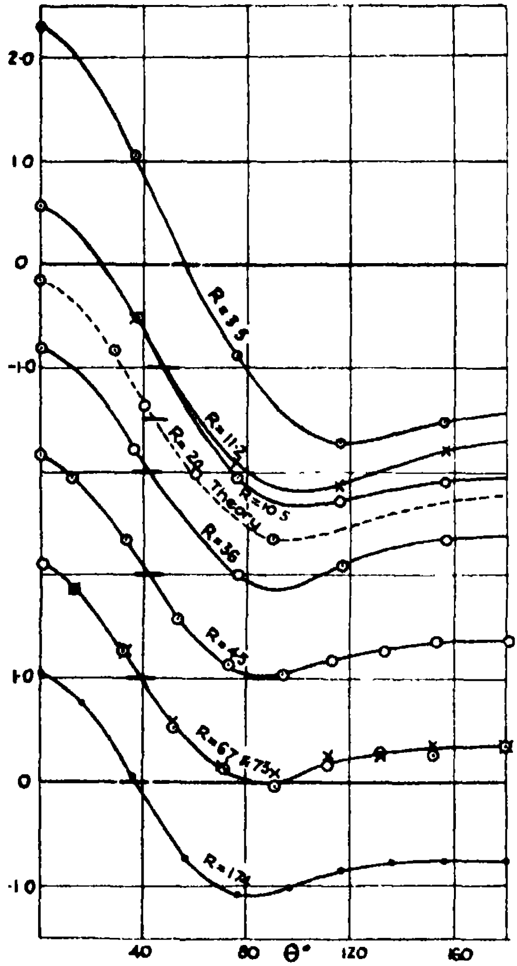


FIG. 6.—Pressure distribution round circular cylinders, experiments in oil and water. Ordinates are $(p - p_0) \div \frac{1}{2}\rho V^2$. The short thick black lines show zero (*i.e.*, static pressure for the curves which they cut).

---

* *Loc. cit.*, R. & M., No. 1194; 'Aero. Res. Ctee.,' R. & M., No. 1389.

Section 2.—*Experimental Determination of the Pressure Distribution Round Circular Cylinders at Low Reynolds' Numbers in Oil and Water.*

The apparatus used was substantially that described in R. & M , No. 1389, and consists of a tank or channel having a working section 5 inches by 5 inches, and a form of Chattock gauge. At low speeds with small cylinders oil is more troublesome than water. This is partly on account of the greater coefficient of expansion of oil at the temperatures used. All leads have to be protected from temperature variation as pressure variations are produced by the slightest change in temperature such as that caused by a hand held near a part of the system. The reason is that the expanding oil is throttled by the very small hole or holes in the cylinder. To appreciate this it should be realized that the pressures to be measured are sometimes only a few thousandths of an inch of water. As the response of the gauge is very slow it takes a considerable time (up to an hour in some cases) to take a single reading of pressure.

The velocity was obtained from the measurements themselves by using the relation given in R. & M., No. 1389 between velocity and front generator pressure. Particulars of the oil used will also be found in that paper.

The static hole was on the channel wall opposite the cylinder, but even so a correction for wall effect is required. The exact amount of this connection is uncertain and the value used was, in the absence of better information, that given in the Appendix to R. & M., No. 1194. This uncertainty makes the measurements made with the smaller cylinders more reliable (other things being equal).

The final results are given in Table I which also contains an explanation giving the method of correction. The procedure of correcting the angles by $\frac{1}{2}h/d$ to allow for the size of the hole is fully justified at higher speeds but not yet at these low values of R (h = hole diameter) The results are plotted in fig. 6.

The values of the coefficient of pressure drag deduced from the values in Table I, are given in Table II, and will be found plotted in fig. 8 along with those from experiments in air (R. & M., No. 1194) and recent results by Linke* who having measured the total drag of cylinders under conditions identical to those obtaining when he measured the pressures, obtains reliable values for the part of the drag due to skin friction. His results bear out the expression given by the writer for the skin friction drag of a cylinder, namely, $K_D = 2 \div \sqrt{R}$†.

---

* 'Phys. Z.,' vol. 32, p. 900 (1931).

† 'Aero. Res. Ctee.,' R. & M., No. 1176 (1928).

Table I.

Oil. R = 3·5.

$t = 15{\cdot}0^{\circ}$ $\nu = 0{\cdot}400$
$d = 0{\cdot}318$ $h = 0{\cdot}045$
$\epsilon = 0{\cdot}005$ $V = 4{\cdot}4$

θ.	$\theta - \Delta\theta$.	p'.
0	0	2·3
40	36	1·05
80	76	−0·88
120	116	−1·72
160	156	−1·52

Oil. R = 10·5.

$t = 15{\cdot}0^{\circ}$ $\nu = 0{\cdot}406$
$d = 0{\cdot}96$ $h = 0{\cdot}11$
$\epsilon = 0{\cdot}022$ $V = 4{\cdot}4$

θ.	$\theta - \Delta\theta$.	p'.
0	0	1·56
40	37	0·49
80	77	−1·07
120	117	−1·29
160	157	−1·10

Oil. R = 11·2.

$t = 15{\cdot}7^{\circ}$ $\nu = 0{\cdot}396$
$d = 0{\cdot}318$ $h = 0{\cdot}045$
$\epsilon = 0{\cdot}005$ $V = 13{\cdot}9$

θ	$\theta - \Delta\theta$.	p'.
0	0	1·54
40	36	0·48
80	76	−0·91
120	116	−1·13
160	156	−0·79

Oil. R = 36.

$t = 15^{\circ}$ $\nu = 0{\cdot}410$
$d = 0{\cdot}96$ $h = 0{\cdot}11$
$\epsilon = 0{\cdot}022$ $V = 14{\cdot}5$

θ.	$\theta - \Delta\theta$.	p'.
0	0	1·19
40	37	0·21
80	77	−1·00
120	117	−0·90
160	157	−0·65

Water. R = 45.

$t = 14^{\circ}$ $\nu = 0{\cdot}0118$
$d = 0{\cdot}142$ $h = 0{\cdot}04$
$\epsilon = 0{\cdot}001$ $V = 3{\cdot}75$

θ.	$\theta - \Delta\theta$.	p'.
0	0	1·16
20	12	0·93
40	32	0·35
60	52	0·42
80	72	−0·87
100	92	−0·96
120	112	−0·81
140	132	−0·73
160	152	−0·61
180	180	−0·62

Water. R = 67.

$t = 14^{\circ}$ $\nu = 0{\cdot}0118$
$d = 0{\cdot}142$ $h = 0{\cdot}04$
$\epsilon = 0{\cdot}001$ $V = 5{\cdot}6$

θ.	$\theta - \Delta\theta$.	p'.
0	0	1·11
20	12	0·88
40	32	0·29
60	52	−0·48
80	72	−0·89
100	92	−1·04
120	112	−0·82
140	132	−0·70
160	152	−0·73
180	180	−0·64

Water. R = 73.

$t = 13{\cdot}6^{\circ}$ $\nu = 0{\cdot}0119$
$d = 0{\cdot}142$ $h = 0{\cdot}04$
$\epsilon = 0{\cdot}001$ $V = 6{\cdot}1$

θ.	$\theta - \Delta\theta$.	p'.
0	0	1·10
20	12	0·87
40	32	0·28
60	52	−0·43
80	72	−0·86
100	92	−0·92
120	112	−0·73
140	132	−0·71
160	152	−0·63
180	180	−0·64

Water. R = 174.

$t = 14{\cdot}4^{\circ}$ $\nu = 0{\cdot}0117$
$d = 0{\cdot}318$ $h = 0{\cdot}045$
$\epsilon = 0{\cdot}005$ $V = 6{\cdot}4$

θ.	$\theta - \Delta\theta$.	p'
0	0	1·043
20	16	0·786
40	36	0·057
60	56	−0·713
80	76	−1·085
100	96	−1·002
120	116	−0·840
140	136	−0·769
160	156	−0·761
180	180	−0·739

$R = Vd/\nu$.
t = temperature °C.
ν = kinematic viscosity cm.^{2}/sec.
d = cylinder diameter cm.
h = hole diameter, cm.
r = (channel width) ÷ d.
$\epsilon = 13 \div (30r + r^2)$.
V = fluid velocity, cm./sec.
$p' = p_1(1 - 2\epsilon) \div \tfrac{1}{2}\rho V^2$.
p_1 = observed pressure difference.

$$\Delta\theta = \frac{h}{2d}.$$

$$p' = \frac{p - p_0}{\tfrac{1}{2}\rho V^2}$$

Table II.

Cylinder diameter (cm.).	Fluid.	Fluid Velocity (cm./sec.).	Reynolds' number.	Coefficient of pressure drag.
0·318	Oil	4·4	3·5	1·325
0·96	,,	4·4	10·5	0·90
0·318	,,	13·9	11·2	0·71
0·96	,,	14·5	36·0	0·49
0·142	Water	3·75	45·0	0·46
0·142	,,	5·6	67·0	0·46
0·142	,,	6·1	73·0	0·45
0·318	,,	6·4	174·0	0·48

The present experiments deal, however, with a different range of Reynolds' numbers (R = 3·5 to R = 240). The determinations of drag by Relf and Wieselberger* from force measurements and the difference between the mean of these and the writers' determination of pressure drag, this difference being the skin friction drag, are also shown in fig. 8.

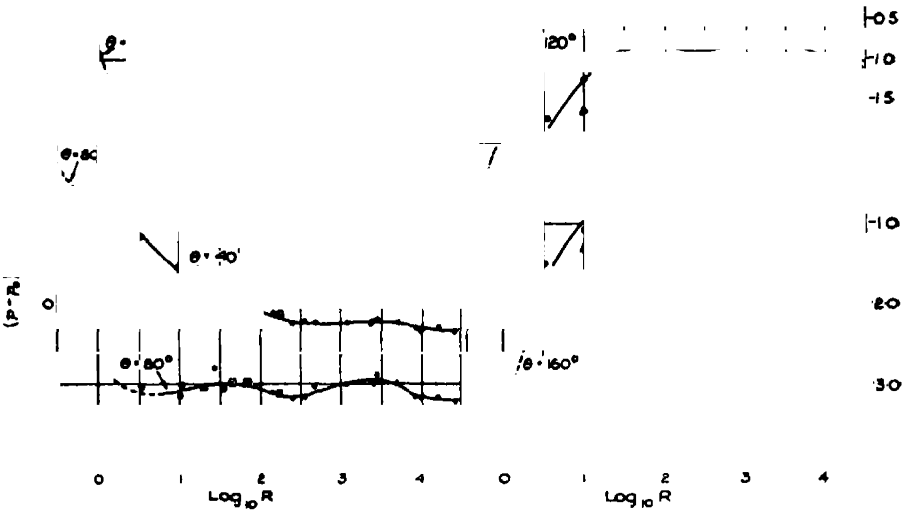


Fig. 7.—Pressure distribution round circular cylinders. △, theory; ○, oil, ⊡, water; ●, air; ▲, air (Linke). Note.—Short dotted curves to right are plotted from Lamb's theory for very low Reynolds' numbers.

Fig. 7 shows the pressures, from the present experiments, from those of R. & M., No. 1194, and from the experiments of Linke. They are plotted for constant values of θ, namely, θ = 40°, 80°, 120° and 160°.

---

* 'Phys. Z.,' vol. 22, p. 321 (1921).

The points shown in triangles are from the arithmetical solutions, and the dotted lines show values obtained from Lamb's analytical solution.* Actually Lamb's solution only applies to values of R less than those shown, in fact to values much less than unity, but evidently in most cases the experimental results are converging with them.

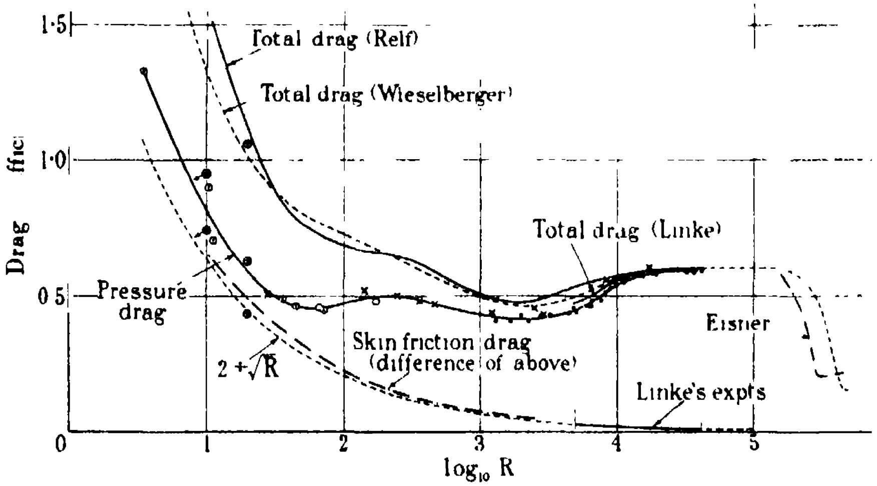


FIG. 8 —Drag coefficients for circular cylinders. ⊙, pressure drag present experiments, oil and water; ×, pressure drag, previous experiments, air; •, pressure drag, Linke's experiments, air; ⊕, arithmetical solutions.

Section 3.—*Photographic Study of the Streamlines at Low Speeds.*

The photographs shown in figs. 9 to 12, Plates 10 to 13, are examples selected from a large number which were obtained by photographing the filament lines produced by introducing streams of coloured water into the flow in the 5 inches by 5 inches tank already illustrated in R. & M., No. 1389.

The best lines were obtained by using fine glass tubes drawn almost to a point but uniform spacing is difficult to obtain in this manner. Accordingly a small brass tube was placed across the channel 2½ inches from the bottom. On the *front* or upstream side of this tube were nine holes 1/64-inch diameter at 1/8-inch spacing from which the ink issued.

The study of the flow by filament lines is limited to low Reynolds' numbers, since it can only be used below the critical speed of the channel. Thus if R_C

---

* Lamb's 'Hydrodynamics,' 4 ed., pp. 581–583.

is the critical Reynolds' number for the channel, and d/l is the ratio of cylinder diameter to channel width, the method is only available up to $R = R_C\, d/l$. R_C was found to be about 2000. Hence if the maximum value of d/l is taken as 0·1, the limiting Reynolds' number for the cylinder is 200. This conclusion is quite independent of the size of channel used.

At very low speeds one of the difficulties met with is the tendency of the ink solution to sink. If the ink density is greater than that of water there is plenty of time for this to occur. For example the speed of the flow for fig. 11*a*, Plate 12, was 0·064 cm./sec. Convection currents are also troublesome.

A curious feature noticeable in figs. 10*a* and 11*a*, Plates 11 and 12, is the tendency of the nine streamlines to close up some distance behind the cylinder to a narrower width than they occupied in the undisturbed stream before they came to the cylinder. This is evidently a three dimensional effect produced by the interference of the walls and bottom. It can hardly indicate an increase in speed but rather a withdrawing of liquid from the plane of the streamlines.

The channel walls have evidently another effect, namely, to delay the production of eddies to a higher speed. An inspection of the two series of photographs, one taken with a ¼-inch cylinder, fig. 10, and one with a ½-inch cylinder, fig. 11, shows this. In the former series the eddies begin about $R = 46$ and in the latter at $R = 62$. An ⅛-inch cylinder was also examined and found to produce eddies when R was but little over 30. The walls evidently prevent the proper formation of the Kármán street with the ½-inch cylinder.

When matters are adjusted so that an ink "streamline" strikes the front of the cylinder and splits, there is still a region close behind the cylinder where the ink does not enter, or at least only a trace enters by sinking slowly along the surface of separation and penetrating at a lower level. As it is difficult or impossible to arrange for such a splitting of a line to take place for any length of time a cylinder was fitted up so that ink could be extruded from a hole in the front generator. The photographs obtained with this arrangement are shown in fig. 12, Plate 13. The supply pipe for the ink is right at the bottom of the channel and the hole on the front generator 2½ inches higher up. The gradual growth of the "kidney" eddy as the speed increases is clearly shown. That this region contains two closed eddies is shown by the difficulty of introducing ink into it and by the ink leaving the surface at right angles. This is further brought out in fig. 12*f* for which photograph the cylinder was turned through 180° so that the ink is emerging from the downstream generator. It travels forward to the "Ablosungspunkt" before it leaves the surface. This photograph has been taken at a higher value of R to show that this phenomenon goes on when the periodic eddies are forming as well as at the lower speeds, showing that the stationary eddies, or at least kindred phenomena, are present along with the periodic eddies.

The frequency of the eddies (f) was obtained by counting the number formed per second. The values of the ratio fd/V so found when plotted on R shows nothing new, the points agreeing well with the values already found by oscillation experiments.†

The tendency for the ratio K/aV to fall with falling Reynolds' number is shown in Table III which contains particulars of the photographs in figs. 10 and 11. The strength of the eddies K must be zero for the Reynolds' number at which the periodic eddies cease to be formed. As already mentioned the periodic eddying behind the $\frac{1}{8}$-inch cylinder shows itself occasionally down to $R = 30$ but it is not very certain below about 36 unless the cylinder is released to oscillate with the appropriate period.

Table III.—Particulars of photographs shown in figs. 10 and 11, Plates 12 and 13.

Photograph.	Cylinder diameter d (cm.).	Velocity V (cm./sec.).	Eddy frequency f	Eddy separation a (cm.)	$R = Vd/\nu$.	fd/V.	K/aV.*
Fig. 10 *a*	0·635	0·067	—	—	3·8	—	—
b	0·635	0·21	—	-	11·0	—	—
c	0·635	0·82	—	—	43·0	—	—
d	0·635	0·93	0·20	4·3	49·0	0·14	0·22
e	0·635	1·02	0·21	4·35	54·0	0·13	0·29
f	0·635	1·97	0·41	3·8	103·0	0·13	0·59
Fig. 11 *a*	1·27	0·064	—	—	7·2	—	—
b	1·27	0·22	—	—	23·0	—	—
c	1·27	0·474	—	—	50·0	—	—
d	1·27	0·634	0·077	—	71·0	0·15	—
e	1·27	0·85	0·11	—	88·0	0·17	—
f	1·27	2·0	0·27	—	208·0	0·17	—

* $K/aV = (1 - af/V)\sqrt{8}$.

In the previous experiments on oscillating cylinders only very small cylinders were used for the lowest values of R. Repeat experiments using the $\frac{1}{8}$-inch diameter cylinder gave the same results, namely, at $R = 33$ oscillation would take place if the natural frequency was adjusted to give fd/V anything between

---

† R. & M., No. 1373, fig. 7, or 'Phil. Mag.,' fig. 3, vol. 12, p. 493 (1931).

0·097 and 0·15, and at R = 28 no oscillation could be obtained at any period. From R = 29 to about R = 35 it seems better to say that the oscillation causes the eddying rather than the eddying the oscillation.

The author previously assumed* that the eddies produced a circulation which produced the side force necessary to maintain the oscillation. An examination through a low power microscope of the ink filament lines in front of the fixed cylinder while eddies were being given off, showed no periodic wavering such as would be expected with a periodic circulation.† It follows either that the eddies produce no circulation or that the circulation is only present when the cylinder is oscillating.

There would seem to be a possible alternative explanation of the phenomenon of oscillation other than that it is produced entirely by the periodic eddies.

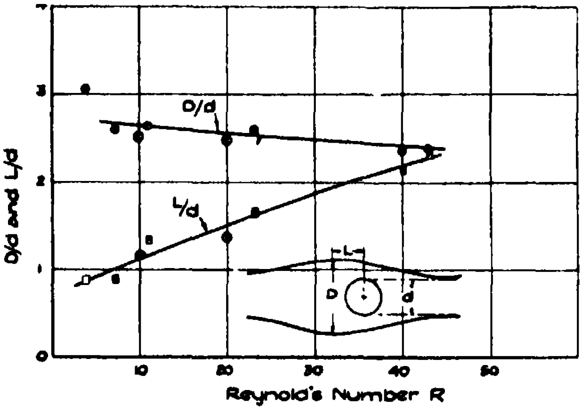


FIG 13.

Is it not possible that the kidney eddies play a part in the following manner? These are, as it were, attached to the cylinder and hence when it is oscillating and approaches the end of its swing they continue, in the same direction long enough to set up a circulation by dragging the liquid round the cylinder with them. The circulation so produced would be of the correct sign to apply the force to drive the cylinder back towards the centre.

To test this hypothesis an $\frac{1}{8}$-inch diameter cylinder was fitted with a vane along its entire length, fig. 14. This vane was free to rotate about the cylinder axis being carried on two needles. The arrangement[*] was found to oscillate at channel speeds above 2·2 cm./sec. (R = 58). For each water speed there was a band of frequencies which gave the maximum amplitude generally corresponding roughly to $fd/V = 0{\cdot}03$ to $0{\cdot}06$. The most interesting point is, however, that the oscillation is very much more vigorous if the cylinder is inclined so that the top is further upstream than the bottom. This was discovered by accident although it might have been expected in view of the explanation given above. Towards the end of the swing the weight of the vane (once it has passed a certain position) is rotating it in the direction necessary for the required circulation. Obviously, however, with the inclined cylinder low water speeds are impossible as the vane falls right over on one side or the other.

---

* R. & M., No. 1373, *loc. cit.*

† Other experimenters claim to have observed this wavering, but probably the Reynolds' number was different.

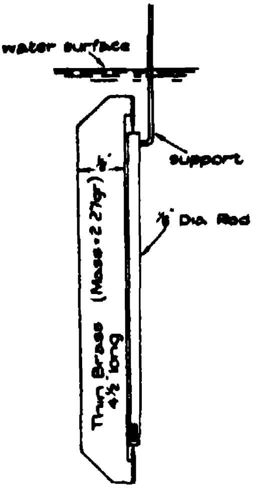


FIG. 14.

The analogy between the vane and the kidney eddies is obviously not perfect, but it is possible that the phenomena are related.

Summary.

In the first part of the paper an arithmetical method of solving the equations of viscous flow is developed. The method is one of successive approximations, whereby the new values are interpolated in such a manner as to involve the fundamental equations being solved. A development of the method involves the use of an auxiliary solution to obtain first a conformal network for the given boundaries. The solution is then carried out on this network. This procedure avoids certain boundary difficulties.

The method is applied to obtain the flow past a circular cylinder at Reynolds' number 20. Two weak symmetrical eddies are obtained behind the cylinder. Values for the drag and pressure distribution are also given.

In the second part of the paper measurements of the pressure round cylinders at very low Reynolds' numbers are given. These are shown to agree with the theoretical values of the first part. A comparison with the measured total drag obtained by other experimenters enables the skin friction drag to be determined.

---

[*] The whole unit was free to oscillate across the channel, as in the experiments described in R. & M. No. 1373, *loc. cit.*

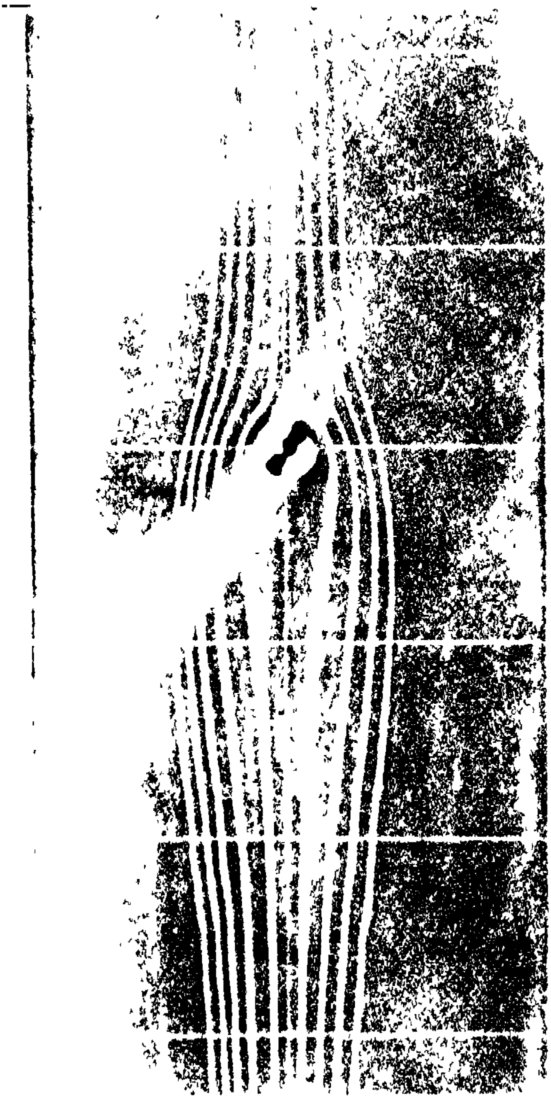


FIG. 9.—½ in cylinder in a 5-in channel R = Vd/ν = 23

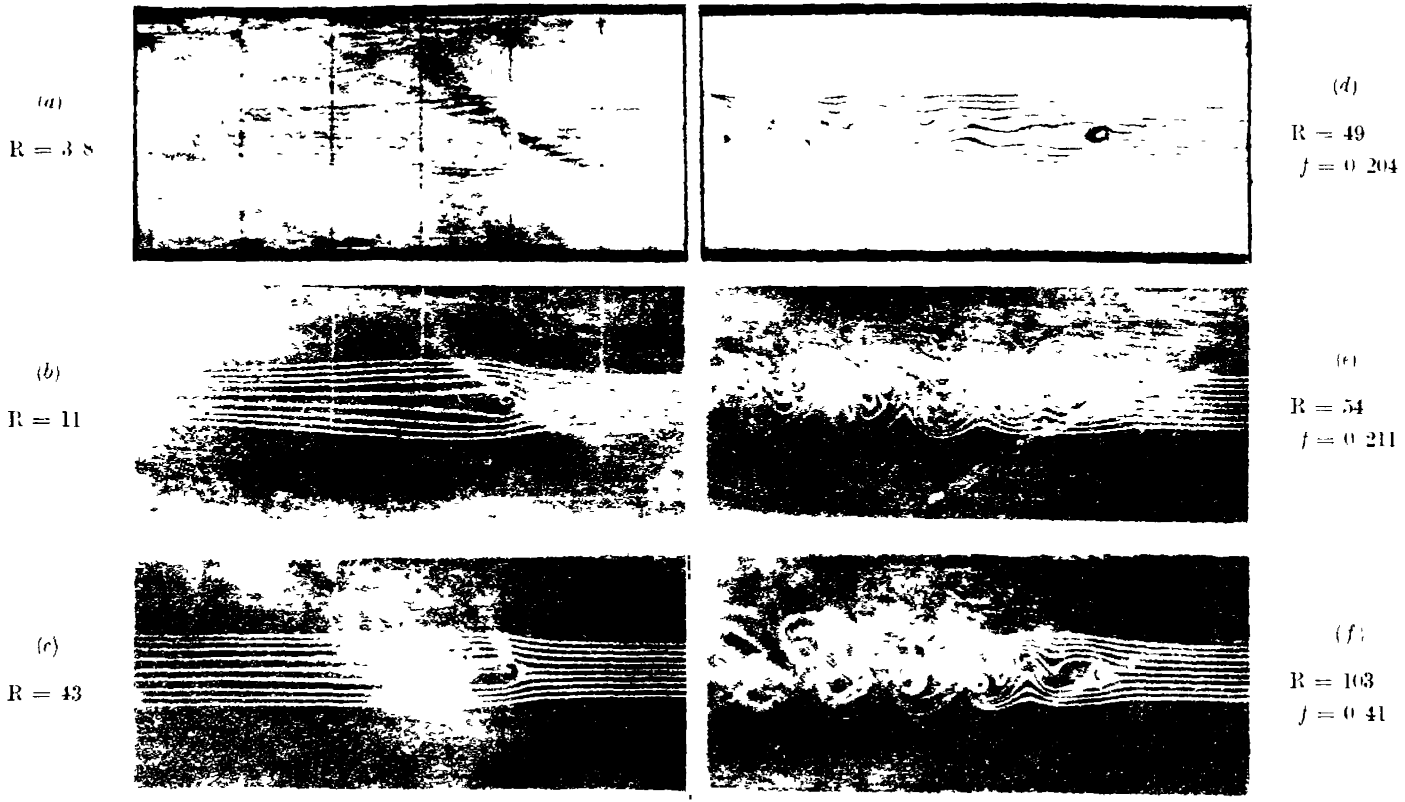
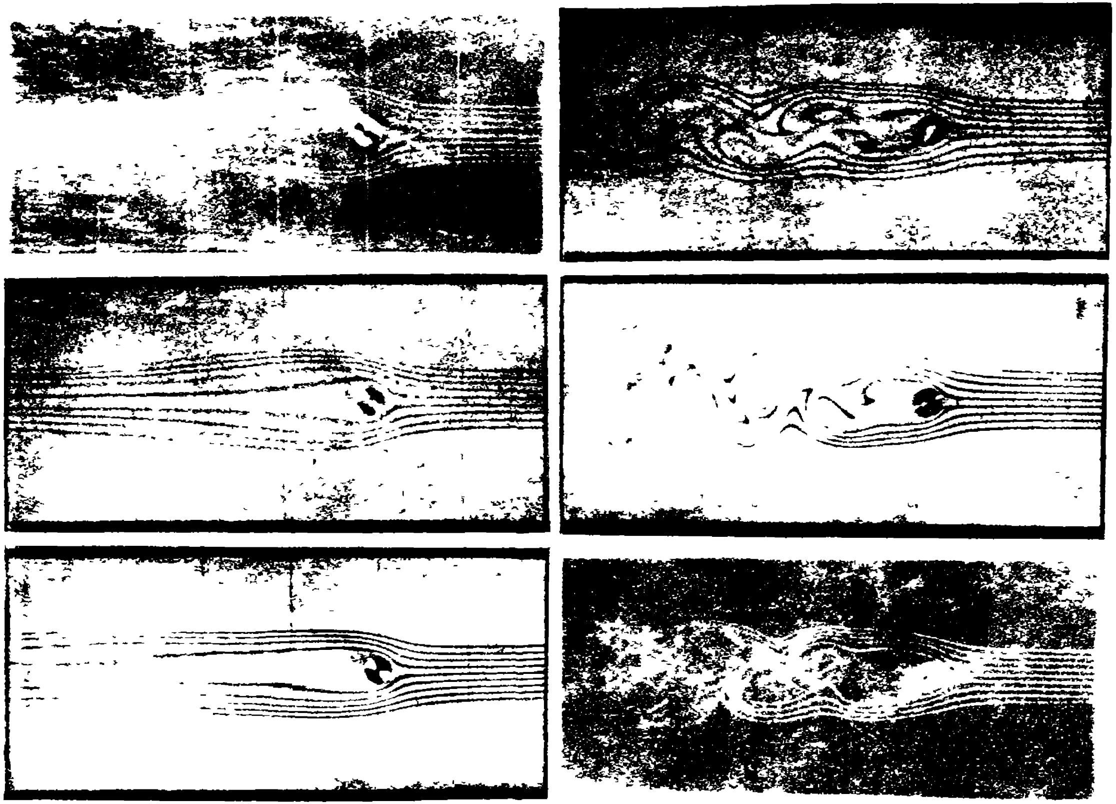


FIG. 10—¼-in cylinder in a 5 in channel



(a) R = 7.2

(b) R = 23

(c) R = 50

(d) R = 71

(e) R = 88 f = 0.109

(f) R = 208 f = 0.269

Fig. 11.—½ in. cylinder in a 5-in. channel.

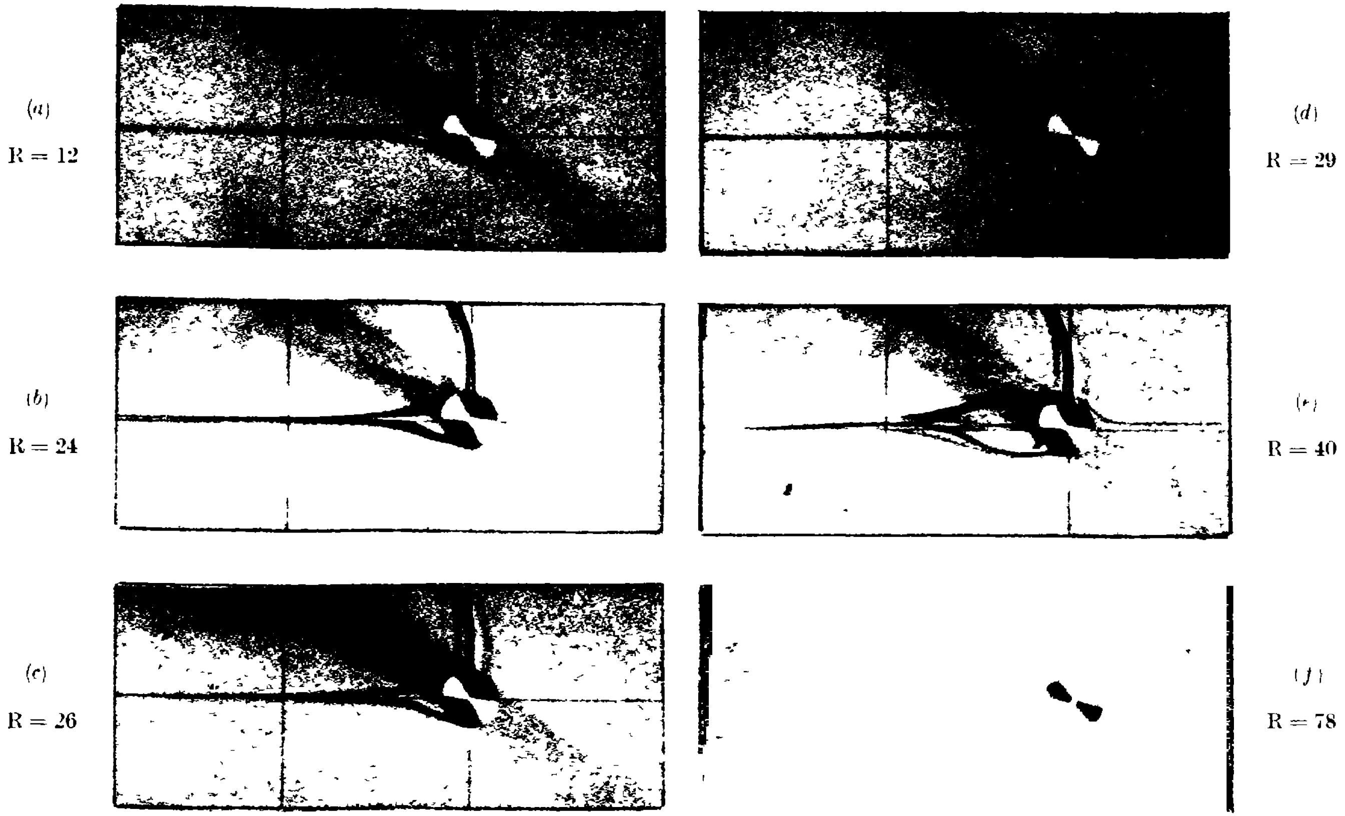


Fig. 12.—½ diameter cylinder.

In the last part a set of photographs of the streamlines past a cylinder is given. A comparison of the theoretical solution with these shows a good agreement. The photographs also show the gradual development and growth of the symmetrical stationary eddies behind the cylinder as the speed increases, followed by the development of the Kármán street. The exact Reynolds' number at which the latter first develops is shown to be affected by the presence of the channel walls.

On the Ionization of Light Gases by X-Rays. I.—*Technique.*

By W. R. Harper, Wills Physics Laboratory, University of Bristol.

(Communicated by Lord Rutherford, O M., F.R.S —Received April 26, 1933)

Introduction.

The measurement of the intensity of an X-ray beam in absolute units is in theory most satisfactorily accomplished by a determination of its heating effect. The method, however, is attended by considerable experimental difficulties, so that its application is very limited, and in practice it is usual to replace it by a determination of the ionization produced when the beam is passed through a gas. To correlate the ionization with an absolute intensity requires a quantitative knowledge of the details of the interaction between the X-rays and the molecules concerned and of the ionization of the gas by the ejected electrons. It sometimes happens that the processes involved about which we know least are relatively unimportant, so that a fairly reliable correlation can be made; and much work has been done on the application of the ionization method to X-ray dosimetry. But in general a quantitative correlation between ionization and intensity is not possible. A further study of the ionization of gases by X-rays is therefore desirable; moreover it may be made to yield important information concerning the processes involved. The early development of the physics of X-rays contains many examples of this, and more recently an important contribution has been made by Stockmeyer.*

The events leading to the ionization of a heavy gas are exceedingly complicated, whereas in the light gases (hydrogen and helium) some of these events are absent or else occur to a negligible extent, so that the interpretation of experiments with the latter becomes simpler and more reliable. These gases are therefore specially worthy of study. Moreover, for them the application of quantum mechanics leads to the most definite results for comparison with experiment, and in particular permits of a direct test of some aspects of Dirac's theory of recoil scattering. The ionization due to the gas itself is, however, very small, and may even be less than the secondary ionization due to electrons liberated from the chamber walls. The technique used in ionization measurements with heavy gases is therefore unsuitable. Hitherto the only attempt made to extend such measurements to light gases is an experiment carried out in 1915 on hydrogen by Shearer* who, however, obtained very variable results and an ionization markedly *smaller* than that to be expected from recoil electrons alone. Moreover his experimental method is now open to criticism in view of our greater knowledge of X-rays, and in particular the fluorescent radiation used was of doubtful homogeneity. The present paper will describe a new technique suitable for quantitative measurements of the ionization produced by X-rays in light gases, and in another paper it will be applied to a re-investigation of hydrogen.

---

* 'Ann. Physik,' vol. 12, p. 71 (1932).

The Pressure Variation Method.

The method adopted by Shearer in correcting for the effect of the electrons liberated from the walls of the ionization chamber is due to Beatty.† It consists in plotting the variation of ionization with pressure up to pressures at which the variation has become linear, when it is assumed that the effect of the walls has saturated, so that the *increase* of ionization is entirely due to absorption of X-rays by the gas. This assumption may, however, be incorrect as is shown by the following experiment carried out by the author. The ionization in a sample of impure hydrogen due to a heterogeneous X-ray beam taken from a glass walled Coolidge tube excited at about 45 kv. was plotted as a function of the pressure. The ionization chamber was cylindrical and of glass, 16 cm. long and 4 cm. in diameter. It was lined throughout with paper made conducting with indian ink, the collecting electrode being an aluminium wire placed to one side of the X-ray beam which passed axially through the chamber. The experimental points are given in fig. 1. Under these conditions there is no possibility of the effect of the walls being saturated as the range* of a 20 kv. electron in hydrogen at N.T.P. is about 7 cm. Nevertheless the ionization pressure curve is a straight line down to a pressure of 150 mm. of Hg, and, moreover, when produced passes very nearly through the origin. There was no difficulty in confirming the result.

---

* 'Phil. Mag.,' vol. 30, p. 644 (1915).

† 'Phil. Mag.,' vol. 20, p. 320 (1910).

A theoretical explanation in general terms can easily be given. The part of the ionization due to the hydrogen and the impurity contained in it would not be proportional to the pressure at low pressures since many of the liberated photoelectrons would be able to reach the walls. It would therefore have a pressure ionization curve of the form of the lower dotted curve in fig. 1, which, however, is not drawn to scale. If the ionization due to the electrons liberated from the end walls by the primary beam were less than that due to those from the corresponding mass of gas, then adding on the ionization they produce would still leave the form of the ionization pressure curve more or less unchanged.† If, on the other hand, the ionization due to the electrons from the end walls were greater than that due to those from the corresponding mass of gas, the ionization pressure curve would have the form shown in the upper dotted curve, provided that the electrons from the gas were unable to reach the side walls. In the general case the curve will be intermediate between the two, and could very well be the straight line found in the experiment.

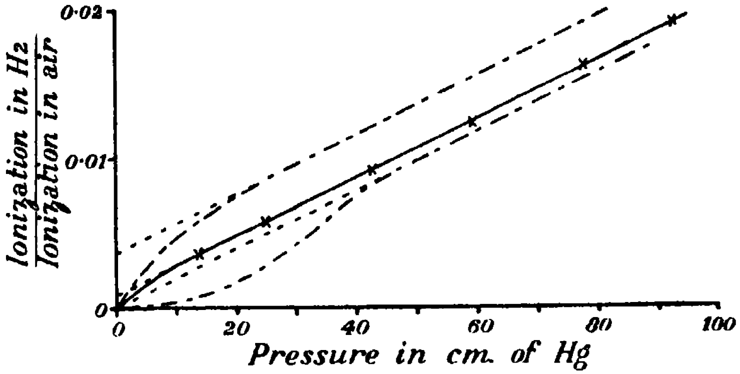


Fig. 1.——Experimental curve ; – – – – theoretical curves ; — — — extrapolations of linear part of curves.

If all the ionization corresponding to the linear part of the curve is attributed to the gas, a comparison with the ionization ratio found for pure hydrogen to air in II (p. 694) makes it necessary to explain the value here found by assuming the presence of far more impurity in the gas than was actually present. The pressure variation method of correcting for the effect of the walls will therefore be rejected, although by going to sufficiently high pressures it would undoubtedly be possible to ensure its reliability.

---

* Nuttall and Williams, 'Phil. Mag.,' vol. 2, p. 1109 (1926).

† W. H. Bragg, "Studies in Radioactivity," chap. XV.

General Considerations.

In Shearer's ionization chamber the wall through which the radiation entered was of mica, and the effect of secondary electrons liberated there was greater than the effect of those liberated in the gas. A simple calculation shows that even with the lightest substance it is feasible to use (namely, celluloid) the effect would certainly not be negligible. The following is the principle of the method used for its elimination in the present work. The rays pass first through a celluloid window into an antechamber in which the secondary electrons from the window are absorbed and the resulting ionization collected; they then pass through a clearance aperture into the main chamber in which the ionization due to the gas is measured, and out of it by a second clearance aperture. Finally they are suitably absorbed at the far end of a second subsidiary chamber similar to the first. In the first form of the apparatus a broad beam of X-radiation was used in order to increase the intensity, and in consequence the clearance apertures subtended a considerable solid angle at the window and at the absorption block. Magnetic fields were therefore applied to coil up the tracks of the unwanted electrons and confine them to the two subsidiary chambers. It was found, however, that the electrostatic fields of the electrodes in the subsidiary chambers penetrated into the main chamber to an extent which enabled them to rob the electrode in the main chamber of a large fraction of the ionization which it should have collected. It was therefore decided not only to use guard electrodes, but also to reduce considerably the size of the clearance apertures. This had the advantage of making it possible to dispense with the magnetic fields. The alternative of making the chamber very large has the inconvenience of involving the preparation of large quantities of pure gas, and moreover loses much of its effectiveness owing to the divergence of the X-ray beam necessitating an increase in the size of the exit clearance aperture. The details of construction of the ionization chamber will be deferred until later when a general consideration of all the processes responsible for ionization in an ionization chamber will be given, and it will be shown that the actual design of the chamber used fulfils the requirement that it must measure the ionization arising from the direct action of the primary beam on the gas, but no secondary ionization such as that from the walls.

Radiation of wave-length about ½ A. is most suitable for the measurement of the ionization from recoil electrons in hydrogen described in II (*loc. cit.*), and the most convenient measure of its intensity is the corresponding ionization in air. The measurement then reduces to a comparison of the ionization in the two gases due to the same X-ray beam, which must clearly be carefully monochromatized. Attempts to use crystal reflection for the purpose led to such small intensities that they were abandoned in favour of a modification of the method due to Ross,* in spite of an inherent defect in the method that a difference effect must be measured and the consequent requirement of increased steadiness of the X-ray beam.

The purity of the gas is of paramount importance. One part of mercury vapour in a million of hydrogen has an effect which can easily be measured. The difficulty of ensuring the requisite degree of purity makes it advisable to have some method of estimating its magnitude. In the measurement of the recoil ionization this can be done by carrying out an additional determination of the ionization ratio on the same sample of gas but for a much longer wavelength, as will be explained in II. The method not only enables one to correct for the effect of impurity but also to correct the observed ionization (due recoil electrons) for any photoelectric effect in the light gas itself.

The Source of Radiation.

Several different X-ray tubes were used in the preliminary investigations dealt with in this paper, but need not be described in detail. They were connected directly across the high tension transformer which was fed from a motor-alternator run off batteries when the inconstancy of the supply mains proved troublesome. In this way it was possible to obtain a beam whose intensity only suffered slow changes with time, so that by sandwiching each reading between two control readings the variations in intensity could be corrected for when necessary.

In the final determinations of the ionization ratios for ½ A., however, no variation of intensity greater than 1% could be tolerated even in the course of several hours. The conditions necessary for ensuring this† I have discussed elsewhere. The X-ray tube used was that described in the first paper referred to in the footnote and illustrated in fig. 2 therein. It had, however, only one filament instead of two as shown. The input to the tube was controlled by method I described in the second paper referred to in the footnote, the high tension voltmeter required being one specially designed for the purpose.* In determining the ionization ratio there is an optimum X-ray intensity which gives sufficient ionization in the light gas to be accurately measurable and an ionization in air (the order of 200 times as great) which is not too large to be conveniently measurable by the method described later.† A current of 20 milliamps at a voltage of 45 kv. D.C. (from a valve rectifier) was found to give a suitable intensity from the copper anticathode. The ripple on the D.C. at this load was about 10%.

---

* 'J. Opt. Soc.,' vol. 16, p. 433 (1928).

† Harper, 'Proc. Camb. Phil. Soc.,' vol. 28, p. 497 (1932); 'J. Sci. Inst.,' vol. 10, p. 10 (1933).

The determinations of the ionization ratios for the longer wave-length were carried out less accurately. For this purpose the tube was excited at 8 kv. peak, and a tube current of 25 milliamps obtained by making the filament sheath 360 volts positive with respect to the filament. Since the same smoothing condensers were used on the valve rectifier their effect under these conditions would be to give only a slight smoothing. Moreover, the high tension voltage was controlled by means of a spark gap, so that the intensity of the X-ray beam could only be *guaranteed* to within 20%, which, however, is sufficient not to impair the accuracy of the measurement described in II. The short wave-length limit of the radiation excited under these conditions is 1·5 A.

The Method of Monochromatization.

Ross (*loc. cit.*) has devised a method of spectral analysis which may equally well be used for monochromatization. Briefly it consists in preparing two filters whose absorption is the same for all wave-lengths except for those in a narrow band between the K edges of the two filters. If then the effect of any heterogeneous X-ray beam is measured first with one filter in the beam, and then with the other, the *difference* between the two effects is entirely due to radiation in the narrow band of wave-length. It is clear from fig. 2 in Ross's paper that the balancing of the two filters for the wave-lengths outside the narrow band can be carried out exceedingly accurately. The only objection to the method as described by Ross is that it is necessary to work with a difference effect of only a few per cent. of the total effect which is actually measured, whereby the accuracy of the measurement is enormously reduced. However, by the additional use of selective filters I have found it possible to reduce the relative intensity of the wave-lengths outside the useful band and thereby increase the difference effect to greater than a quarter of the effect with the element of lower atomic number in the beam The loss of accuracy in taking the difference is then not nearly so serious. The loss of intensity due to the selective filters is considerable, but the available intensity is in any case so large that it was still found possible to obtain the intensity required.

---

* Harper, 'J. Sci. Inst.,' vol. 10, p. 13 (1933).

† The use of a gas heavier than air would have been very inconvenient because of the excessive ionization.

The balanced filters used were of silver and cadmium which give a wave-length band from 0·485 A. to 0·463 A., these being also the filters that were specially tested by Ross. The metals used were of the highest purity obtainable, and were rolled* to thicknesses of 0·045 mm. (Ag), and 0·047 mm. (Cd). They were mounted in the manner described by Ross, the final adjustment for thickness being carried out by altering the tilt. The uniformity of absorption over the working area is of great importance and was therefore tested, using a narrow X-ray beam, and shown to be better than 1%. The foils were balanced on a spectrometer with wide slits using the MoK lines (0·67 A.). Care was taken that the slit width was sufficiently small to exclude radiation in the range of wave-lengths between the K edges of the filters, and that no appreciable radiation from the unreflected beam could enter the ionization chamber either directly or by scattering. It was unnecessary to exclude second order reflection. After balancing, the mean of seven ionization readings for silver differed from the mean of seven for cadmium by only one part in 600.

When using the balanced filters the maximum difference effect is obtained when the maximum of the continuous spectrum from the tube lies between the K edges of the filters, which for the pair used occurs for about 50 kv. on the tube. The voltage actually used was 45 kv. and was therefore not far from the optimum. This has the additional advantage of giving a nearly uniform wave-length distribution between the K edges. The selective filters used to increase the difference effect were of tin and strontium bromide. Their uniformity is of no great importance. They resulted in a reduction of intensity of the order of 50 times but increased the difference effect from 4% to 27% (referred to the effect with silver in the beam), *i.e.*, by a factor of about 7.

As a check on the working of the method it was used to determine the half value thickness for the absorption in aluminium. In fig. 2 is plotted the log of the ionization against the number of foils of aluminium. The upper curve is for Cd in the beam, the middle one for Ag in the beam, and the lower one for the difference effect, *i.e.*, for a wave-length band between 0·485 A. and 0·463 A. The two upper curves both show a marked systematic deviation from linearity showing that the radiation was not homogeneous, but the points for the difference effect lie on a straight line within the limits of experimental error which, however, are greater for the difference curve. The half value thickness calculated from the curve is 1·5(7) mm. The mean half value thickness for the above band of wave-lengths* is 1·5(0) mm. The agreement is therefore very satisfactory.

---

* I am indebted to Messrs. Johnson, Matthey for carrying out the difficult operation of rolling these foils.

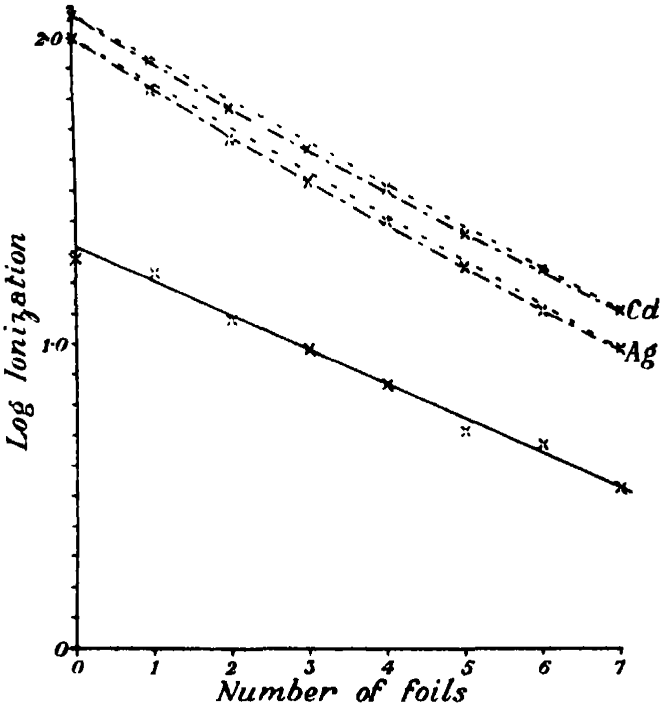


FIG. 2. ——— experimental curves; —— difference curve; ---- straight lines.

The Ionization Chamber.

A diagram of the ionization chamber and its connections is given in fig. 3. The main chamber A is a hollow cylinder of cast gunmetal, one end being part of the casting, the other being screwed directly on to a flange and the join made tight by painting the edge with cellulose enamel. The cylindrical brass

* From Kirchner, "Allgemeine Physik der Röntgenstrahlen."

antechamber B and the similar end chamber C are soldered to the ends of the main chamber eccentrically as shown. D and D′ are the two circular apertures (in lead screens) which limit the X-ray beam, D being placed close to the window of the X-ray tube. Between D and D′ there is a clearance aperture E in a lead screen which prevents stray radiation from reaching the second limiting aperture. The celluloid window at F is rather larger than the aperture D′. G and G′ are clearance apertures. The apertures external to the chamber are carried by a brass tube and arm H bolted to the end plate, and J is an ebonite disc serving to insulate D from the chamber. The balanced foils are situated at K and the selective filters at E. The whole of the main chamber and the end of the chamber C are lined with thick aluminium which in turn is lined with indian inked paper L. M and M′ are electrodes of aluminium wire which collect the ionization from the two subsidiary chambers, N and N′ are the guard electrodes, and P the collecting electrode. The electrodes in the main chamber are of indian inked paper, rectangular, and 3·4 cm. broad. The collecting electrode is insulated by pyrex glass protected by an earthed guard ring Q. The ionization chamber is connected to the gas apparatus at R.

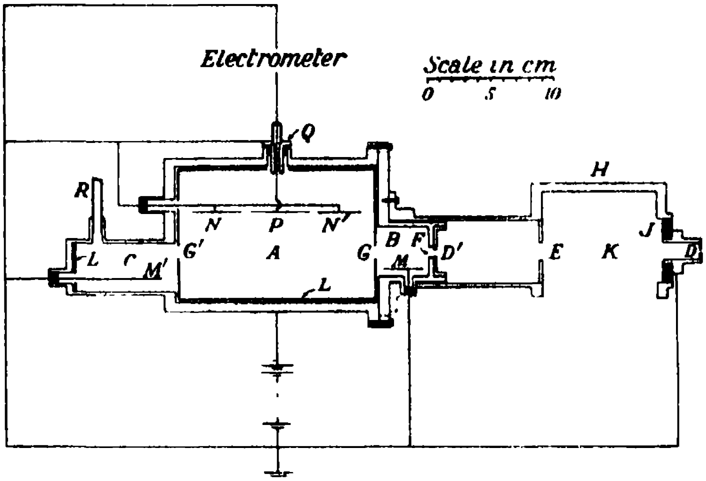


Fig. 3.

The Domain from which the Ionization is Collected.

When a narrow beam of X-rays passes through a gas of small absorption coefficient* and indefinite extent the ionization produced occupies a cylindrical column with the beam as axis, and the ionization is uniform along the column. What is required to be measured in comparing two gases is the ratio of the ionization between two parallel mathematical planes perpendicular to the axis for one gas and the ionization between the same two planes for the other gas. Firstly, therefore, it is necessary that the side walls of the chamber and the collecting electrode shall be outside the cylindrical column of ionization. At atmospheric pressure in air the range* of a photoelectron liberated by a quant of the harder radiation used is about $1\frac{1}{2}$ cm., the maximum range in hydrogen of a recoil electron from such a quant is less than one-tenth of this, and the range in hydrogen of a photoelectron liberated by a quant of the softer radiation used is about 1 cm. Referring to fig. 3 it is seen that the clearances in question are more than twice the requisite amount. Photoelectrons liberated in hydrogen by the harder rays, however, will be able to reach the walls, but although in air the majority of the ionization is caused by photoelectrons, in hydrogen recoil electrons are of primary importance, and a correction can be applied for the incomplete absorption of such photoelectrons as are present in hydrogen. Secondly, since some of the ionization in the domain from which it is collected is due to electrons starting outside that domain, the end walls must not interfere with such electrons. They must therefore be distant from the domain by an amount greater than the range of an electron. This is ensured by the guard electrodes. Thirdly, since the ionization in hydrogen is confined to a much narrower column than in air, it is necessary for the collecting domain to have plane ends, though since the ionization is symmetrically distributed about the axis these planes need not be accurately perpendicular to the axis. Some method is, however, required for mapping the lines of force in the chamber, since even in the presence of the guard electrodes they may be appreciably curved.

---

* A small correction may be required for the absorption of the beam in the gas before it reaches the collecting domain if this is different for the two gases compared, and for the change of intensity of the beam during its passage through the collecting domain.

If a two-dimensional scale model of the chamber is constructed this can be done either by the use of gypsum crystals to show up the direction of the lines† or by adapting the model as an electrolytic cell on A.C. and exploring the lines of force with probe electrodes and a pair of phones.‡ Neither method is, however, capable of giving quantitative results since both simplify the three-dimensional problem to a two-dimensional one. The following method was therefore adopted. A suitably constructed α-particle gun was fitted into the antechamber B and a narrow beam of α-particles was fired through the aperture G into the main chamber. A polonium source was used and to prevent contamination of the apparatus it was totally enclosed in the gun and the α-particles allowed to escape through a very thin cellophane window. The residual range* of the particle was measured, and after correcting for the stopping in the cellophane agreed to within ½ mm. with the range as measured by G. I. Harper and Salaman.† When the ionization chamber was filled with air at atmospheric pressure the α-particle beam did not reach the domain of the collecting electrode P. The pressure was then reduced until a sudden commencement of the ionization current to P denoted the arrival of the α-particles in the collecting domain. Then, knowing the pressure at which this occurred and the direction of the beam, a point on the end surface of the collecting domain could be calculated, and by varying the direction of the beam the end surface could be mapped out.

---

* Nuttall and Williams, *loc. cit.*

† Stockmeyer, *loc. cit.*

‡ Luhr and Bradbury, 'Phys. Rev.,' vol. 37, p. 998 (1931).

The main chamber being symmetrical it was only necessary to carry out the measurements for one end. Five points on the end surface of the collecting domain were determined, one on the axis of the X-ray beam, and four more 2 cm. from it and equally spaced round it. They were all coplanar within the limit of experimental error which was about 1 mm Their plane was, however, inclined at about 8° to the end wall of the chamber.

The Robbing Effect.

In an ionization chamber constructed for preliminary tests having no guard electrodes and the clearance apertures G and G′ relatively much larger, it was found that about 20% of the ionization formed in the main chamber was collected by the electrodes in the subsidiary chambers. The extent of this robbing effect made it advisable to verify experimentally that in the chamber of fig. 3 only the domain from which the guard electrodes would otherwise collect was robbed.‡ The lines of force were investigated by the electrolytic cell method, and it was found that for the larger aperture the domain of the subsidiary chamber electrode extended ⅔ cm. into the main chamber. This gives a factor of safety of at least 6, sufficient to cover the error introduced by working with a two-dimensional model. It should, perhaps, be mentioned that both in an electrolytic cell on A.C. and in an ionization chamber which all but the largest currents are being measured, space charges are negligible so that the lines of force are as *in vacuo.*

---

* Extrapolated ranges were used throughout.

† 'Proc. Roy. Soc.,' A, vol. 127, p. 175 (1930).

‡ The test described in the last section did not make this unnecessary, as it was carried out with the subsidiary chamber electrode removed.

The ionization in air due to secondary electrons from a paper window is very small compared with the direct ionization of the gas by the rays. If then a beam of X-rays is passed through the chamber and the ionization collected by P measured, this will be uninfluenced by the insertion of a thin conducting paper window at G, unless the field of M penetrates sufficiently far into the main chamber to affect the domain of P. The experiment was performed using an arrangement whereby the window could be moved into and out of the beam from the outside of the chamber, and the change did not amount to 1%.

Processes responsible for Ionization in an Ionization Chamber.

When a beam of X-radiation passes through an ionization chamber all processes resulting in the formation of secondary electrons which can be absorbed by the gas contribute to the ionization in the chamber. The following is a *complete* list of all the theoretically possible processes :—

(1) The action of the primary beam on the gas-giving photoelectrons, recoil electrons, and electrons due to the Auger effect.

(2) The action of the primary beam liberating secondary electrons from the windows.*

(3) The absorption by the gas of radiation from the gas—
   (*a*) Scattered.
   (*b*) Fluorescent.

(4) The absorption by the gas of radiation from the windows—
   (*a*) Scattered.
   (*b*) Fluorescent.

(5) The liberation of secondary electrons from the walls due to the absorption of—
   (*a*) Scattered radiation from the gas.
   (*b*) Fluorescent radiation from the gas.
   (*c*) Scattered radiation from the windows.
   (*d*) Fluorescent radiation from the windows.

---

* The term " window " refers to any part of the wall through which the primary X-ray beam passes, and the term " walls " is to be taken as including windows.

(6) The absorption by the gas of radiation from the walls—
  - (*a*) Scattered.
  - (*b*) Fluorescent.

(7) The liberation of secondary electrons from the walls by radiation from other parts of the walls—
  - (*a*) Scattered.
  - (*b*) Fluorescent

In general it is required to measure (1) + (3*b*), the ionization arising from the secondary emission from the gas due to the action of the primary beam. The necessary conditions for measuring (1) have already been discussed. The fluorescent radiation from air is extremely soft and will be completely absorbed in a chamber of the dimensions of the one shown in fig. 3, but hydrogen and helium are peculiar, and will be considered in II.

It is required *not to measure anything but* (1) + (3*b*). (3*a*) is a second-order effect for a chamber like the one in fig. 3, and is therefore negligible. (5) (5*a*) (5*b*) (6) and (7) will collectively be termed the " wall effect," and (2) (4) (5*c*) and (5*d*) the " window effect." The method used for eliminating the window effect has already been indicated, and full details will shortly be given.

*The Wall Effect.**

(5*b*) is small because (3*b*) is nearly complete. (7*a*) and (7*b*) are second-order effects relative to (5*a*) which will be shown to be negligible. (6*a*) is a second-order effect relative to (3*a*) which is negligible. It remains to consider (5*a*) and (6*b*).

The thick aluminium lining of the gunmetal chamber reduces any fluorescent radiation from the gunmetal to a negligible amount, and any fluorescent radiation from the aluminium is reduced to a negligible amount by the paper, the fluorescent radiation from which can be neglected. This disposes of (6*b*). The relative magnitude of (1) and (5*a*) depends on the details of the absorption and scattering of the radiation by the gas, and will therefore be considered in II where (5*a*) will be shown to be negligible even for hydrogen.

---

* In constructing the chamber it is convenient to leave small screws, etc., exposed, and it is important that these should not appreciably increase the wall effect. This was verified to be so by an approximate direct experimental determination of the effect of completely lining the inside of the main chamber with a heavy metal (tin).

The Window Effect.

The details of the arrangement for eliminating this are shown in fig. 3, and in fig. 4 which shows the aperture system to scale (different in the two directions). The solid angle subtended by the aperture G at F is less than $0\cdot005 \times 2\pi$, and the solid angle subtended by G′ at the far end of C is also less than $0\cdot005 \times 2\pi$. As there is no reason to expect any *great* preponderance of emission of window electrons in the forward direction, and as they will be partially absorbed in the subsidiary chambers, the ionization they produce in the main chamber will be cut down by the aperture system to less than one hundredth of its full value, and this will be greatest in hydrogen. In the chamber of fig. 3 the effect on the ionization in hydrogen of inserting a thin conducting paper window at G was determined. The experiment measured directly that part of the window effect which reaches the domain from which P collects. It was one-tenth of the ionization due to (1) + (3*b*) + (5*a*). The aperture system therefore reduces the window effect to less than a thousandth of the main effect.

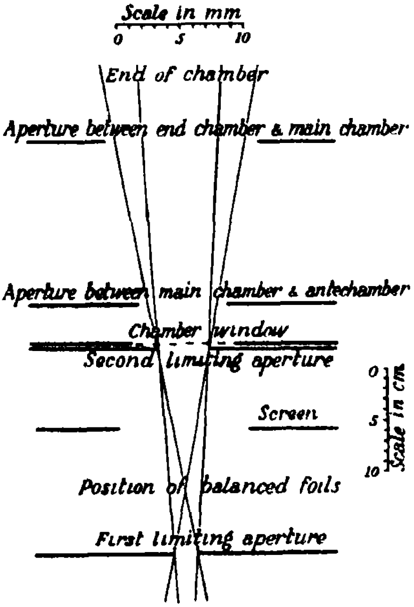


FIG 4.

The Aperture System.

The ionization chamber and aperture system could be bolted together to ensure the correct alignment of the apertures, and the whole was very rigidly constructed. It could be taken to pieces and put together again without disturbing the alignment. The two limiting apertures D and D′ and the clearance aperture G could be taken off as a rigid whole, and the fact that no rays could fall on the edge of G verified by eye. G′ was tested in the same way before the end chamber was soldered on. With the chamber completed, however, it was again tested photographically for clearance and alignment.

The Measurement of the Saturation Currents.

To minimize any ionization due to stray radiation the lead from the ionization chamber was only 15 cm. long, and the Compton electrometer was shielded by a lead housing. The earthing key was attached to the ionization chamber so that only two insulators were required in the electrometer lead, the one in the ionization chamber and the one in the electrometer itself.

The only satisfactory way of comparing two ionization currents which differ greatly in magnitude is to neutralize them by means of a charge induced on the collecting system through a condenser, using the electrometer as a null instrument. The ratio of the currents is then given in terms of the condenser voltages and the times if the capacity is the same for both. The condenser was incorporated in the electrometer itself. The necessary voltage was obtained from a high tension accumulator battery of about 60 volts when carrying out measurements on air, and connected direct across a high resistance potentiometer. For hydrogen the high-resistance potentiometer was connected across a low-resistance potentiometer which was connected across a 2-volt accumulator. in this way a fraction of a volt could be obtained. In all cases the full drop across the high-resistance potentiometer was used. The reduction of voltage in the low-resistance potentiometer was calculated from the resistances which were checked against a standard box. The voltages of the two batteries were read on a calibrated voltmeter.

When measuring the ionization current in hydrogen a sensitivity of about 3000 mm. per volt at 1 metre was used, and the rate of drift was about 5 mm. in 8 seconds. (With soft rays and the samples of hydrogen on which the measurements described in II were carried out there was a drift of about 2 mm. in 7 seconds.) The method of taking the readings was to start the needle from well to one side of the zero by applying a suitable potential to one of the pairs of quadrants via the earthing key, and start the stop-watch when the spot of light passed the zero. Half way through the observation the full charge was induced on the collecting system, and subsequently the time of repassing the zero taken. In this way the effect of a small leak is eliminated, though the leak was quite negligible. The method ensures that the rate of drift has become steady at both instants when the time is taken.

When measuring the ionization current in air the sensitivity was reduced by earthing the needle. Under these conditions the deflection is proportional to the square of the voltage on the quadrants. It was about 40 mm. for 1 volt. A rather different procedure was adopted in taking the readings as owing to the rapidly changing sensitivity the drift is never steady. The needle was started from its zero and timed on passing a given deflection. It was then immediately brought back to the zero by the potentiometer, and kept there until the full charge had been induced on the collecting system. This could be done very easily and accurately as the needle is very insensitive to changes of quadrant voltage at the zero. The needle was then allowed to drift and timed on passing the same deflection as before. Now before both the beginning and end times the needle drifted in exactly the same way so that the lag in its drift was automatically allowed for. The spot of light passed the position of timing at 3 to 4 mm. per second. (This method does not eliminate the effect of a leak, but it reduces it.)

All "natural effects," apart from the residual ionization in the chamber, were less than 1% of the effect measured. The electrometer system was tested with a γ-ray source and found to give readings consistent to within ± 1%. The residual ionization in the chamber due to radioactive contamination and cosmic rays was of course quite negligible compared with the ionization in air, but was sufficient to cause irregularities in the readings for hydrogen. It was verified that the irregularities in the readings with the X-rays on could be quantitatively accounted for by the occasional bursts of ionization that occurred with the rays off, and that the effect disappeared when the chamber was evacuated. The bursts of ionization were undoubtedly due to α-particle contamination. The total residual ionization in hydrogen was about $2\frac{1}{2}$ pairs of ions per cubic centimetre per second at atmospheric pressure. (In air it was about 15.) This residual effect is no larger than is to be expected from unavoidable contamination, and any attempt to eliminate the α-particles would have led to awkward complications. Since, however, the irregularities in readings of 2 to 3 minutes were almost always less than ± 3%, and were undoubtedly subject to a Gaussian distribution law, a mean of 10 readings should be correct to within ± 1%. A set of 10 readings was therefore taken for each measurement of all ionization currents.*

The Test of Saturation.

The lack of saturation in an ionization chamber in which the ionization is due to X-rays and of the order of magnitude here measured has been discussed by Webster and Yeatman,† and shown to be due almost entirely to initial recombination when a moderate voltage is applied to the electrodes. The precise shape of the saturation curve therefore depends on the precise initial distribution of the ions, so that it may be different for different gases and for different radiations. For a gas ionized by X-rays there is no theory which enables this shape to be calculated, so that one cannot say on theoretical grounds what the lack of saturation will be under given conditions. An accurate experimental determination of the shape of the saturation curve is attended by considerable difficulties and leads to a curve the slow rise of which makes it very difficult to determine accurately the asymptote, particularly as there is no definite theoretical indication as to the shape of the extrapolated part of the curve, and as the lack of saturation that is required to be estimated comes entirely from this extrapolated part. At the best, then, estimating the lack of saturation is a very uncertain procedure. In the present work it was verified that halving the working voltage on the chamber did not decrease the ionization current by as much as 1%, 240 volts being required for hydrogen and 360 volts for air. Any lack of saturation should therefore have been very small, and since the minimum voltage consistent with the above condition was used for both gases the lack of saturation should have been about the same for both gases so that its effect on the ionization ratio must have been quite negligible.

---

* With hard rays in hydrogen the residual ionization is eliminated when measuring the difference effect, but with soft rays it has to be measured and allowed for.

† 'J. Opt. Soc.,' vol. 17, p. 254 (1928).

Acknowledgment.

Originally the work was undertaken at the Cavendish Laboratory at the suggestion of Professor Lord Rutherford, to whom I am indebted for his interest and encouragement. I also wish to thank Dr. J. Chadwick for much valuable advice, and to acknowledge a grant from the Department of Scientific and Industrial Research which enabled me to commence the research. It has been completed at the Wills Laboratory at Bristol and I am greatly indebted to Professor A. M. Tyndall for his interest in the work during its continuation here.

Summary.

A technique is described for investigating the ionization produced in hydrogen and in helium by X-rays uniformly distributed in the wave-length band 0·485 A. to 0·463 A., and by softer rays of wave-length about 1½ A. The processes responsibe for ionization in an ionization chamber are analysed in detail and it is shown that the method described measures in general · (I) the ionization due to the secondary electrons ejected by the primary beam from the gas ; (II) the ionization due to reabsorption by the gas of the fluorescent radiation excited by the primary beam in the gas ; (III) the ionization due

to the liberation of secondary electrons from the walls of the chamber by the radiation scattered by the gas; but that it eliminates all ionization due to other causes. In II (*loc. cit.*) (III) will be shown to be negligible. The standard of X-ray intensity adopted is the corresponding ionization in air, so that the measurement reduces to a determination of the ionization ratio for two gases and the same X-ray beam.

On the Ionization of Light Gases by X-Rays. II.—The Ionization of Hydrogen by Recoil Electrons.

By W. R. Harper, Wills Physics Laboratory, University of Bristol.

(Communicated by Lord Rutherford, O.M., F.R.S.—Received April 26, 1933.)

Introduction.

In Part I, p. 669, a technique has been described for determining the ratio of the ionization in a light gas (hydrogen or helium) to that in air when ionized by the same X-ray beam, homogeneous rays of medium wave-length and soft heterogeneous rays being available for the measurement. The ionization ratio can be converted into the ratio of the energies absorbed by the two gases by making use of the known value of the ratio of the energies required to form a pair of ions in the two gases; and since the ionization in air is due almost entirely to photoelectrons and the absorption coefficient is known, the energy in the incident beam can be obtained from the energy absorbed by air; thus the energy absorbed by the light gas can be correlated with the energy in the incident beam. The radiation of medium wave-length (about ½ A.) ionizes the light gas chiefly through the agency of recoil electrons, so that after applying a correction (obtained from the soft ray ratio) for the ionization due to photoelectrons, the fraction of the energy in the incident beam converted into recoil electron energy by the gas may be obtained, and compared with the predictions of the quantum theory of recoil scattering. In this paper the comparison is carried out with measurements on hydrogen, and for convenience it will be made between the experimental and calculated values of the ionization ratios.

Excluding the early work of Shearer already mentioned in I no experimental determination of the total energy associated with recoil electrons has hitherto been made by any method as direct as the present one, though less direct methods have been employed.* In general, however, the experimental technique was open to criticism, and the interpretation of the measurements uncertain, so that it is not surprising that the results were inconsistent either with one another or with theory. It is also possible to calculate the total energy associated with recoil electrons from other experimental facts concerning recoil scattering, but the experimental errors involved combine to make the final result very unreliable. The energy associated with recoil electrons is, however, not only of theoretical interest but also of great practical importance, since all effects arising from scattering, in the scattering substance itself, are due to recoil electron emission.

Calculation of the Ionization Ratio.

The theory of the recoil scattering of X-rays by an atomic system has not yet been rigidly developed on the basis of quantum mechanics, but a simplified treatment has led to some fairly certain results. Wentzel† has shown that the distribution of recoil electrons from a hydrogen atom, and presumably therefore from a hydrogen molecule, is the same as from free electrons if the binding energy is small compared with the recoil energy, and if $\alpha = h\nu/mc^2$ is small, where the symbols have their usual significance. These two conditions are fulfilled for the vast majority of the recoil electrons from hydrogen for a wavelength of ½ A. Moreover, Wentzel concludes that in all probability when the theory is properly applied the conclusion will be that the total intensity of modified scattering will be the same as for free electrons. I shall therefore assume that the hydrogen gas may be taken as equivalent to a distribution of two free electrons per molecule when calculating the recoil scattering, so that unmodified scattering is negligible. There is abundant experimental evidence‡ that unmodified scattering may be neglected for light elements and radiation of ½ A.

The theory of the scattering of X-radiation by free electrons has been fully developed.§ Formulæ are given for the fraction of the incident energy removed from the primary beam by the scattering process, for the distribution of the intensity of the scattered radiation with scattering angle, and for the ratio of the energy of a recoil electron to that of the scattered quant as a function of scattering angle. From this may be calculated the fraction of the energy of the primary beam that is converted into recoil electron energy.

---

* Fricke u. Glasser, 'Z. Physik,' vol. 29, p. 374 (1924); Fricke, 'Nature,' vol. 116, p. 430 (1925); Schocken, 'Z. Physik,' vol. 64, p. 458 (1930).

† 'Z. Physik,' vol. 58, p. 348 (1929).

‡ Kirchner, "Allgemeine Physik der Röntgenstrahlen," p. 453.

§ Dirac, 'Proc. Roy. Soc.,' A, vol. 111, p. 405 (1926); 'Proc. Camb. Phil. Soc.,' vol. 23, p. 500 (1926); Gordon, 'Z. Physik,' vol. 40, p. 117 (1927); Klein and Nishina, 'Z. Physik,' vol. 52, p. 853 (1929).

Assuming a uniform distribution of energy in the wave-length band from 0·463 A. to 0·485 A., the mean value of $\alpha = h\nu/mc^2$ is found to be 0·0511 using the values of h, m and c given by Birge.* α^2 is therefore negligible with respect to unity. If θ is the angle through which a quant is scattered, then the ratio $f(\theta)$ of the energy of the recoil electron to the energy of the scattered quant is $\alpha(1 - \cos\theta)$. Let $\mathrm{I}(\theta)$ be the dependence of the intensity of the scattered radiation on scattering angle, and F the ratio of the total energy associated with recoil electrons to the total energy of the scattered radiation. Then

$$\mathrm{F} = \frac{\int \mathrm{I}(\theta) f(\theta)\, d\omega}{\int \mathrm{I}(\theta)\, d\omega},$$

where $d\omega$ is an element of solid angle, and the integration extends over the whole range of 4π.

Now† $$\mathrm{I}(\theta) = (1 + \cos^2\theta)\{1 + \alpha(1 - \cos\theta)\}^{-3},$$

whence

$$\mathrm{F} = \frac{\alpha\int_0^{\pi}(1 + \cos^2\theta)(1 - \cos\theta)\{1 + \alpha(1 - \cos\theta)\}^{-3}\sin\theta\, d\theta}{\int_0^{\pi}(1 + \cos^2\theta)\{1 + \alpha(1 - \cos\theta)\}^{-3}\sin\theta\, d\theta}.$$

Substituting $1 - \cos\theta = y$ we have

$$\mathrm{F} = \frac{\alpha\int_0^{2}(2 - 2y + y^2)(1 + \alpha y)^{-3}\, y\, dy}{\int_0^{2}(2 - 2y + y^2)(1 + \alpha y)^{-3}\, dy}.$$

Expanding by the binomial theorem with neglect of a remainder that may easily be calculated to be less than 1%,

$$\mathrm{F} = \frac{\alpha\int_0^{2}(2 - 2y + y^2)(1 - 3\alpha y + 6\alpha^2 y^2)\, y\, dy}{\int_0^{2}(2 - 2y + y^2)(1 - 3\alpha y + 6\alpha^2 y^2)\, dy}.$$

---

* 'Phys. Rev. Supplement' (1929).

† Dirac, *loc. cit.*, equation (40).

Evaluating the integrals and rejecting all terms contributing less than 1% this becomes

$$F = \frac{\alpha(5 - 21\alpha + 66\alpha^2)}{5 - 15\alpha + 42\alpha^2}.$$

Hence the ratio of the energy associated with recoil electrons to the energy removed from the primary beam $= F/(1 + F) = 0\cdot0461$, using the calculated value of α.

Now the energy removed from the primary beam per second is*

$$I_0 \frac{2\pi N e^4}{m^2 c^4} \frac{1+\alpha}{\alpha^2} \left\{ \frac{2(1+\alpha)}{1+2\alpha} - \frac{1}{\alpha} \log(1 + 2\alpha) \right\},$$

where I_0 is the intensity of the primary beam in ergs cm.$^{-2}$ sec.$^{-1}$, N is the total number of electrons and where the other symbols have their usual significance. Evaluating this to 1% with the aid of seven figure logarithms it becomes

$$1\cdot212\ I_0\ 2\pi N e^4/m^2c^4.$$

It follows that the energy given per second to recoil electrons is

$$\begin{aligned} &0\cdot0559\ I_0\ 2\pi N e^4/m^2c^4 \\ &= 0\cdot0559\ I_0\ 4\pi e^4/m^2c^4 \text{ per molecule} \\ &= 5\cdot52\ .\ 10^{-26}\ I_0 \text{ cm.}^2\text{/molecule} \\ &= 1\cdot49\ .\ 10^{-6}\ I_0 \text{ cm.}^2\text{/c.c. of gas at N.T.P.} \end{aligned}$$

Consider now the ionization in air. If τ is the true absorption coefficient, and σ' the scattering absorption coefficient, the energy given per second to photoelectrons and recoil electrons is

$$(\tau + \sigma')I_0 \text{ cm}^3\text{/c.c. of gas.}$$

The total absorption coefficient μ is equal to $\tau + \sigma + \sigma'$, where σ is the true scattering coefficient. The above expression may therefore be written as

$$(\mu/\rho - \sigma/\rho)\,\rho I_0 \text{ cm.}^3\text{/c.c. of gas,}$$

where ρ is the density. There has been no direct determination of μ/ρ for the requisite wave-length of 0·474 A., but Schocken† has measured it in the range 0·12 A. to 0·42 A. The value of μ/ρ obtained by extrapolating his results to 0·474 A. using formula (5) of his paper is 0·436 cm.^{2} . gm.$^{-1}$. There is no entirely satisfactory experimental determination of σ/ρ, but there is abundant experimental evidence that for radiation of wave-length 0·474 A. in air the scattering will be chiefly modified,* so that the formulæ already applied to hydrogen will also apply approximately to air. The equations are in this case only required to determine a small part of the ionization, since the recoil ionization in air is very much smaller than the photo-ionization, and to determine the magnitude of σ/ρ which is here not very different on the quantum theory assuming all the scattering to be modified and on the classical theory assuming it all to be unmodified. Furthermore, the theoretical values of σ/ρ agree very well with the experimental result of Barkla and Sadler.† No great error will therefore be incurred by assuming that all the scattered radiation is modified, so that the value of σ/ρ deduced from the quantum theory may be used.

---

* Dirac, *loc. cit.*, p. 423.

† 'Z. Physik,' vol. 58, p. 39 (1929).

We have, therefore,

$$\frac{\sigma + \sigma'}{\rho} = \frac{\sigma_0}{\rho} \cdot \tfrac{3}{4} \cdot \frac{1 + \alpha}{\alpha^2} \left\{ \frac{2(1 + \alpha)}{1 + 2\alpha} - \frac{1}{\alpha} \log (1 + 2\alpha) \right\}$$

$$= 0 \cdot 909\, \sigma_0/\rho,$$

where σ_0 is the classical value.

Since $\sigma'/\sigma = F = 0 \cdot 0483$, we have $\sigma/\rho = 0 \cdot 867\, \sigma_0/\rho = 0 \cdot 173$ gm.$^{-1}$ cm.^{2}.

ρ for dry air at N.T.P. is $1 \cdot 29 \,.\, 10^{-3}$ gm./c.c. Using these figures the energy given per second to photoelectrons and recoil electrons is found to be

$$3 \cdot 41 \times 10^{-4}\, I_0 \text{ cm.}^2\text{/c.c. of gas at N.T.P.}$$

Referring back to the hydrogen calculations we have for the ratio of the energy given to recoil electrons in hydrogen to the energy given to both recoil and photoelectrons in air the value 0·00437, whence the ratio of the ionization due to recoil electrons in hydrogen to the ionization due to recoil and photoelectrons in air is 0·00437 times the ratio of the volts per ion pair for air to the volts per ion pair for hydrogen.

The only published value of this ratio is that of Lehmann,‡ who found a constant value of 1·22 for electrons of energies in the range 200–600 electron volts. Now the ratio is required for electrons of energy about 1 electron kilovolt in hydrogen, and about 25 electron kilovolts in air. Lehmann, however, concludes that the volts per ion pair for air is constant up to 1 electron kilovolt, and Buchmann* and Eisl† reach the same conclusion working in the ranges 4–13 and 10–60 electron kilovolts respectively, though it must be admitted that the values obtained by Lehmann and by Buchmann and Eisl disagree by a factor of 1·4. Nevertheless, in the present unsatisfactory state of the subject it seems reasonable to suppose that the ratio is constant so that Lehmann's value may be used. In using it we assume that the ionization collected in the experiment is caused by the ejected electrons but includes neither the ejected electron itself nor its parent positive ion, since this is what corresponds to Lehmann's conditions of measurement. For air the correction is quite negligible, but for hydrogen and helium it would be worth applying (being about 3%) were it not for the uncertainty already attaching to the value. Moreover, about half the error is compensated for by the fact that the fluorescent radiation from the gas is not reabsorbed and its energy must be added to the energy of the recoiling electron to obtain the energy lost by the X-ray quant.

---

* Woo, 'Phys. Rev.,' vol. 28, p. 426 (1926).

† 'Phil. Mag.,' vol. 17, p. 739 (1909).

‡ 'Proc. Roy. Soc.,' A, vol. 115, p. 624 (1927).

We therefore have finally for the *calculated ratio of the ionization in hydrogen due to recoil electrons to the ionization in air, for an X-ray beam of wave-length uniformly distributed in the band* 0·485 *A. to* 0·463 *A., the value* 0·0053. The uncertainty in this value arises chiefly from that attaching to the volts per ion pair ratio. For softer rays it follows from the formulæ that the ratio is roughly inversely proportional to the fourth power of the wave-length.

The calculated ionization ratio differs from the ratio that is measured by the apparatus described in I in that it does not include either ionization due to photoelectric effect in the light gas, or ionization due to secondary electrons liberated from the walls of the chamber by radiation scattered by the gas. The experimental ratio will be corrected for the former, and the latter will now be shown to be negligible.

The Wall Effect.

An upper limit for the ionization due to this cause ((5*a*) in the notation of I) may easily be obtained by comparing that part of the scattered energy which is reabsorbed in a surface layer of the walls of thickness equal to the range of a secondary electron, with the energy absorbed by the gas. The change of wave-length on scattering may be neglected in an approximate calculation. For a wave-length of about ½ A. in hydrogen it was found in the last section that 20 times as much energy is scattered as is absorbed. The range of the corresponding secondary electrons in carbon* is about 10^{-3} cm. A layer of carbon of this thickness absorbs a fraction of the incident X-ray energy equal to $7 \cdot 10^{-4}$. The ratio of the energy absorbed in this layer to the energy absorbed by the gas is therefore $1 \cdot 4 \cdot 10^{-2}$. The actual magnitude of (5*a*) must be several times smaller than this upper limit, so that (5*a*) is negligible for ½ A. radiation in hydrogen (and similarly for helium). For longer wave-lengths the thickness of the surface layer from which the secondary electrons can escape is inversely proportional to the square of the wave-length, so that assuming the absorption coefficient to be proportional to the cube of the wave-length the fraction of the incident energy absorbed by the layer is proportional to the wave-length. The ratio of the energy scattered to the energy absorbed (by scattering) by the gas is also roughly proportional to the wave-length, so that the upper limit for (5*a*) varies as the square of the wave-length. At 1½ A. therefore (5*a*) is less than 0·13 of the ionization from recoil electrons. Reference to the final results shows that this latter ionization was only a small part of the measured total ionization, which again makes (5*a*) negligible in view of the smaller accuracy of the soft ray determination.†

---

* 'Ann. Physik,' vol. 87, p. 509 (1928).

† 'Ann. Physik,' vol. 3, p. 277 (1929).

Purification of the Gases.

The air used was purified from CO_2 by means of caustic potash, and dried over P_2O_5.

The hydrogen was obtained from a cylinder and contained only about 0·1% impurity, this being chiefly oyxgen. It was admitted via a small storage bulb through a glass enclosed needle valve into the apparatus, and then passed over spongy copper gauze heated to 420° C. to convert the oxygen into water, and subsequently over P_2O_5 to remove the water. The gas was then passed through activated charcoal immersed in liquid air to remove other impurities, and a liquid air trap placed next to the ionization chamber served to eliminate mercury vapour. Before the purification the whole apparatus was pumped out with a Hyvac and tested for leaks, and it was verified that any impurity from residual gas or leaks could not have affected the hard ray measurements by as much as 1% The quantity of gas to be purified was about 1½ litres, and the purification generally took about 1½ hours. Altering the time to 1 hour or 3 hours did not, however, affect the results.

---

* Nuttall and Williams, 'Phil. Mag.,' vol. 2, p. 1109 (1926).

† 1½ A. is the lower limit of wave-length of the heterogeneous beam used, so that the upper limit for (5*a*) is underestimated; but this may be overlooked since the upper limit must be a considerable overestimate of the true value.

The pressure of the gas was measured by means of a mercury manometer, and its temperature by a thermometer placed outside the chamber. The ionization measurements were reduced to N.T.P. on the assumption that the ionization was proportional to the density of the gas.

Variation of the Ionization with Pressure.

The correction mentioned in the footnote on p. 677 of I arising from the absorption of the beam in the gas in the ionization chamber is easily seen to be negligible except when using the soft rays in air. It was eliminated here also by working at a pressure of about 200 mm. which has the additional advantage of reducing the ionization current to a manageable amount. The ionization was found to be proportional to the pressure at about this pressure, but at higher pressures fell off owing to the absorption correction, fig. 1. The ionization due to the soft rays in hydrogen was too small to permit of a satisfactory investigation of the ionization pressure curve. The ionization due to the hard rays in air was proportional to the pressure down to ½ atmosphere, fig. 1. With hard rays in hydrogen, however, the arrangement for eliminating the window electrons cannot be expected to work efficiently at pressures much below 1 atmosphere. Apart from this the ionization should be proportional to the pressure even with heterogeneous rays (which permit measurements of much

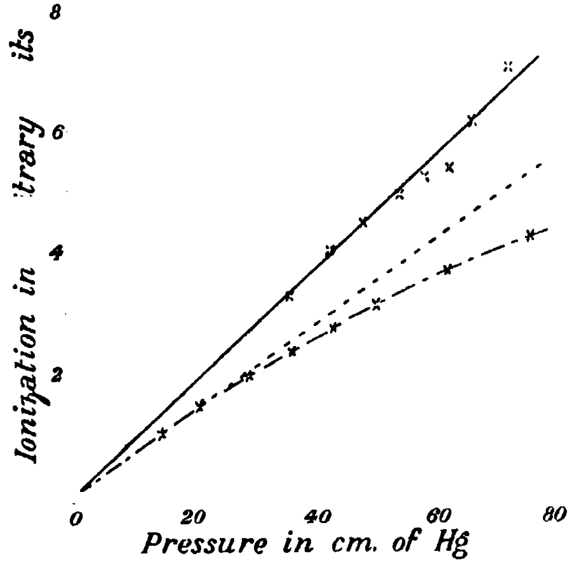


FIG. 1. ——— hard rays ; – – – – soft rays.

greater accuracy). On reducing the pressure from atmospheric to 453 mm. of mercury the ionization (with either filter in the beam) was 3% greater than it would have been if the pressure ionization curve had been linear. The discrepancy is to be attributed to window electrons, but arises entirely from the measurement at the lower pressure, and therefore throws no doubt on the validity of the measurements taken at the higher pressure.

Experimental Procedure and Results.

The procedure in carrying out a determination of the ionization ratios was as follows. A sample of hydrogen was purified (the final pressure being about 1 atmosphere), and a measurement made of the ionization due to the soft radiation. The conditions of working the X-ray tube were then altered and a measurement made with the hard radiation. The hydrogen was then replaced by air at atmospheric pressure, and a measurement of the ionization in this gas made with the hard radiation. The pressure was then reduced, the X-ray tube altered to give the soft radiation as before, and a soft radiation measurement made on air. After the hydrogen had been purified the measurements took about 6 hours, but as already mentioned the constancy of the X-ray beam could be guaranteed for even longer periods. Determinations were carried out on four different samples of gas, and the results are given in the following table.

Sample.	Soft ray ratio.	Hard ray ratio.
I	0·00032	0·0058
II	0·00020	0·0053
III	0 00020	0·0053
IV	0·00023	0·0052
Mean	0·00023	0·0053 (4)

The results obtained with sample I are weighted ½ in taking the mean because the individual observations were rather less consistent than stated in I, owing to difficulty being experienced in keeping the tube current and voltage sufficiently steady. Remembering that the determination of the hard ray ionization currents in hydrogen is subject to an experimental error of about 1% in the mean, the error being considerably less for air, discrepancies of up to about 10% are to be expected in the hard ray ratio which depends on a 20% difference between two measured currents. The consistency of the four results is therefore satisfactory. Since by far the largest error in the whole measurement comes from the necessity of taking a difference, the final result should be correct to about 5%, it being a mean of four determinations.

The Soft Ray Ratio.

Since the ratio of the ionization in hydrogen due to recoil electrons to the ionization in air varies roughly inversely as the fourth power of the wave length, this ratio for the soft rays should be at most about 0·00005, and is therefore only able to account for about a fifth of the observed soft ray effect. The rest must be due either to photoelectric effect from the hydrogen or from impurity. Without further experiment it is not possible to decide which, but the fact that (with the exception of sample I) the soft ray ratios are consistent to within the rather large experimental error of about 20% suggests that the photoelectric effect in hydrogen itself may account for a large part of the remaining ionization. The soft ray ionization ratio does, however, give an upper limit for the atomic absorption coefficient of hydrogen in terms of that of air. If again Lehmann's value of the volts per ion pair ratio be assumed we have:

Ratio of the atomic absorption coefficient of hydrogen to the atomic absorption coefficient of air $< 0{\cdot}0002$ *for heterogeneous rays of wave-length rather greater than* $1\frac{1}{2}$ *A.*

It is of interest to note that the rough empirical law that the atomic absorption coefficient is proportional to the fourth power of the atomic number leads to a value of 0·0004. The law therefore breaks down for hydrogen.

The Hard Ray Ratio.

The hard ray ratio as experimentally determined has to be corrected so as to exclude any photoelectric effect. Now whether the observed photoelectric effect with soft rays is due to hydrogen itself or to impurity, that part of the ionization ratio which is due to photoelectrons will not differ very much with soft and hard rays unless an absorption edge intervenes, *i.e.*, unless the impurity has a high atomic number, and any such impurity even if present would certainly have been removed by the liquid air and charcoal.* Neglecting the contribution of the recoil electrons to the soft ray ionization, the required correction could therefore be made by merely subtracting the soft ray ratio from the hard, were it not for the fact that the photoelectrons from the hard rays in hydrogen are not completely absorbed in the chamber. A fraction only of the soft ray ratio has therefore to be subtracted. This fraction was determined by introducing a considerable impurity (about 1% of air) into a sample of hydrogen, and measuring both soft and hard ray ratios. Subtracting the values found with pure hydrogen gave the effect of the impurity alone, from which it was found that the fraction in question was 0·8. Applying the correction to the hard ray ratio we obtain :

---

* The possibility that mercury vapour had not been completely removed was ruled out by carrying out a set of four determinations on hydrogen purified in the usual way but saturated with mercury vapour. Making reasonable assumptions it is possible to calculate approximately how much the ratios should have been affected by the presence of the mercury, and the calculated differences agreed with the observed to within the limits of experimental error.

Experimental ratio of the ionization in hydrogen due to recoil electrons to the ionization in air, for an X-ray beam of wave-length uniformly distributed in the band 0·485 *A.* *to* 0·463 *A.* = 0·0052. The result should be correct to about 5%.

The calculated ratio was 0·0053, subject to an error of uncertain amount.

Summary.

The technique described in I (p. 669) is applied to a determination of the magnitude of the ionization in hydrogen produced by recoil electrons liberated by homogeneous X-rays. Agreement is obtained with the value calculated using Dirac's theory of recoil scattering as a basis. An upper limit is also obtained for the photoelectric absorption by hydrogen which is considerably lower than the value obtained from the empirical law that the atomic absorption coefficient is proportional to the fourth power of the atomic number.

Meeting for Discussion on the Ionosphere.

(June 22, 1933)

Professor E. V. APPLETON, F.R.S.: In opening this meeting it is of interest to recall that a discussion on the same subject was opened by Sir Ernest, now Lord, Rutherford, before the Society about seven years ago. I, for one, have found it very interesting and instructive to read again the account of that meeting in the Proceedings. I noticed, for example, that we are now able to report substantial progress during the intervening seven years, so that it is now possible to answer in greater detail and with more certainty some of the questions which were raised at that discussion. It is also gratifying to note that various suggestions were then made which have turned out to be very fruitful and particularly helpful to the experimental workers in this field.

In the first place, it is worth noting that our subject has acquired a name. The last discussion was on " The Electrification of the Upper Atmosphere," whereas our present discussion is on " The Ionosphere," a term first suggested by Mr. Watson Watt. In opening the discussion I feel that, perhaps, it ought to be my task to report progress on the various lines mentioned by Sir Ernest Rutherford and, in doing so, to go very generally over the whole field, leaving the detail to be filled in by later speakers. I propose also to deal chiefly with information which has been obtained by the methods in which I have been specially interested, namely, those of radio exploration.

In the absence of data derived from measurements *in situ*, such as are possible for the lower strata of the atmosphere, information concerning the nature of the ionosphere is derived from ground observations on (1) terrestrial magnetism; (2) luminous manifestations such as the auroral electric discharges, meteorites, etc.; and (3) wireless wave exploration. Although the first indication of pronounced upper-atmospheric electrification came from (1), the prosecution of (3) has proved, on the whole, the most fruitful. Wireless methods possess a marked advantage in that an exploration can be made at any time and it is not necessary to wait for natural sequences or irregularities.

Wireless exploration consists in projecting waves (usually) vertically upwards and noting the characteristics of the waves reflected from the ionosphere. The quantities measurable are (*a*) the group-time of flight on the up and down journey, (*b*) the polarization, and (*c*) the intensity of the returned waves. By multiplying the group-time by the velocity of light we obtain the equivalent path P′ which is greater than the actual path traversed by the waves. On the other hand, the optical path P is less than the actual path, so that where P can be estimated (as is possible, by an extrapolation, in some cases) upper and lower limits for the actual height of reflection can be given. To measure (*a*), some distinguishing characteristic must be impressed on the up-going wave such as a change of frequency or a change of amplitude. Both methods have been used but the amplitude-modulation method gives the more easily interpretable data. The actual magnitudes involved in the group-time measurements are of the order of a millisecond, and so do not present great instrumental difficulties. Measurements of polarization and intensity become, for the wave-lengths used (*e.g.*, 20 to 2000 metres) merely alternating electric force measurements within an area small compared with the wave-length.

From measurements of (*a*), (*b*) and (*c*) information concerning the nature of the ionosphere has been derived which may be summarized under the following heads :—

(A) *Structure.*—The most direct index of ionospheric conditions is obtained by measuring the equivalent path P′ for a range of frequencies *f*. For certain values of *f* there are found discontinuous increases of P′ which are interpreted as indicating the existence of maxima of ionization. Usually there is one break in the (P′, *f*) curve, indicating the penetration of the lower (Region E) of the two main regions into which the ionosphere is divided. An example of a (P′, *f*) curve for an undisturbed day is shown in fig. 1, where the discontinuity in the ordinary ray curve at 3·1 megacycles/sec. is clearly seen. At this frequency reflection from the lower Region E ceases to be marked and reflection from Region F begins. The noon ionization in the upper Region F is usually 3½ to 4 times that in Region E, average (equinox) values being $6 \cdot 1 \times 10^5$ and $1 \cdot 8 \times 10^5$ electrons per cubic centimetre respectively.* Neither the diurnal nor seasonal variation of Region F is as marked as that of Region E, a difference probably accounted for by the difference in the pressures at the two atmospheric levels in question. There can be little doubt that the maximum of Region E ionization is reached at a height of about 100 km. above the ground. The height of the corresponding Region F maximum is much more difficult to estimate, but a figure of 180 km is probably not greatly in error.

---

* These values have been calculated from an application of the theory of dispersion with the inclusion of the Lorentz-Hartree "polarization" term. If this term is neglected, since its inclusion is still a very debatable matter, the values given should be multiplied by ⅔.

Experiments carried out within the last twelve months at Slough, England, have brought to light further details concerning what may be called the fine-structure of the ionosphere. The intermediate region between Regions E and F has been shown not to be devoid of ionization, and, on relatively infrequent occasions, evidence has been found of another maximum of ionization there. But it is not usual in England at noon for this intermediate region to be as strongly ionized as Region E.

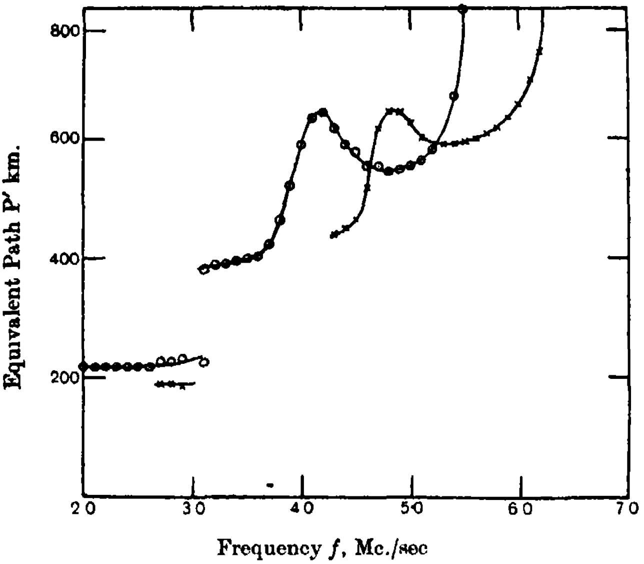


Fig. 1.—(P′, *f*) curve at 1200 G.M.T., Slough, April 10, 1933.
⊙ Ordinary ray ; × extraordinary ray.

Evidence has also been found of the existence of a protuberance or ledge on the underside of Region F, the maximum ionization of which increases with solar altitude. It may not be superfluous to point out that any region higher than Region F, but of lower ionization, would not be detected by the usual wireless methods of exploration.

(B) *Diurnal, Seasonal and other Regular Variations.*—All regions of the ionosphere show a pronounced diurnal variation showing that they are normally influenced by a solar radiation which travels in straight lines. The ionization is replenished daily at a rate dependent on solar altitude and, during the night, it decreases steadily due to the processes mentioned in (F) below. The maximum ionization in Region E is found at local noon while that in Region F occurs most probably one or two hours later (though evidence is conflicting on this latter point). The simple diurnal variation curves on undisturbed days bear a strong resemblance to the theoretical relations worked out by Chapman for the ionizing influence of a monochromatic radiation on a rotating atmosphere. The experimental ratio of summer noon to winter noon ionization is about 2·2 for Region E and, with less certainty, 1·5 to 1·8 for Region F.

Weekly measurements of the maximum noon ionization in Region E made from 1931 onwards, together with other observations, less direct but extending over a longer period, suggest that there is a direct connection between the undisturbed value of noon ionization and the sunspot cycle. Ionization at sunspot maximum appears to be 50% to 60% greater than at sunspot minimum. An important matter yet to be settled is whether this variation is to be attributed to an 11-year cycle in the intensity of ultra-violet light from the sun.

(C) *Recurrence Tendencies and Irregularities.*—Measurements of ionization densities have demonstrated conclusively that the ionization of both Regions E and F can increase at times during the night when it is impossible for solar radiation, travelling rectilinearly, to be effective. Abnormal values for Region E ionization have also been noted, particularly in summer, which are greatly in excess of those expected from the seasonal variation. A (P′, *f*) curve for such an abnormal summer day is shown in fig. 2, which should be compared with fig. 1. In fig. 2, it will be seen that the ionization in Region E is so intense that waves of a frequency as high as 7·0 megacycles/sec. did not penetrate it. Moreover, the very small variations of P′ with *f* indicates an exceedingly sharp gradient of ionization with height. That such conditions were not entirely local is shown by the fact that similar data were obtained on the same day, both at Slough and at the Halley Stewart Laboratory at Hampstead.

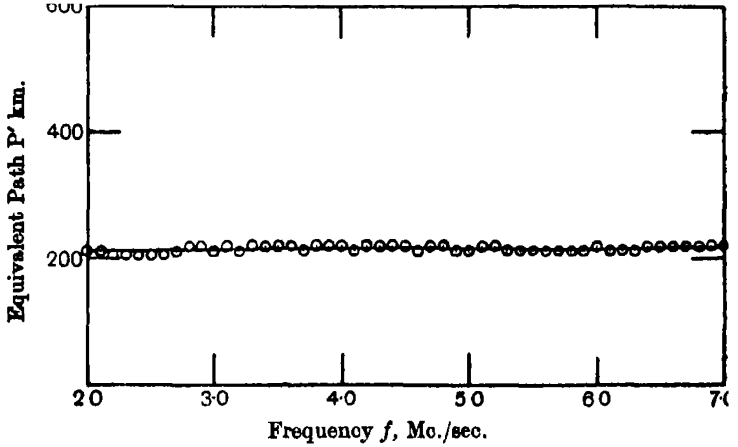


FIG. 2.—(P′, *f*) curve at 1200 G.M.T., Slough, June 16, 1933.

As possible causes of this abnormally high Region E, ionization thunderstorms and high-speed charged particles from the sun have been suggested. Thunderstorms, as C. T. R. Wilson pointed out nearly 10 years ago, might be expected to influence upper-atmospheric ionization, either by "runaway" electrons or by ionization by collision produced at high levels owing to the intense electric fields involved.

Certain types of abnormal nocturnal increases of ionization become more marked the higher the latitude indicating the influence of the earth's magnetic field on the tracks of the charged ionizing particles. Moreover, the correlation of wireless data with sunspots and magnetic storms, and the existence of a 27-day recurrence tendency, further strengthen the well-established view of a pronounced solar control. The usual explanation that the charged particles come from the sun has much in its favour, but there is need for a fuller exploration of the possible influences of thunderstorms in geophysical phenomena. It is not yet certain that the world's thunderstorms do not act, to some extent, as an intermediary mechanism in providing the charged particles which cause phenomena exhibiting the 27-day recurrence tendency and it would be a matter of great interest to examine the thunderstorm data for such a tendency and also for an 11-year period. A first hint in this connection was noted in the measurements of Mr. Naismith and myself on the seasonal variation of Region E ionization at noon. These data which, with Mr. Lutkin's assistance, were found to indicate the correlation of abnormally high summer ionization with thunderstorm activity, also suggested a 26- or 27-day period. In pursuing the matter further, I am much indebted to Dr C. E. P. Brooks of the Meteorological Office who very kindly supplied me with details of the available literature on the subject. The closest relation hitherto found between the frequency of thunderstorms and solar activity is that of Septer,* who examined the data obtained from 229 Siberian stations over the period from 1888 to 1924. The three sunspot maxima and the four minima are faithfully reproduced in the curve of sunspot activity, except that the double spot maximum of 1905 and 1907 is represented by a single thunderstorm maximum in 1906. Dr. Brooks found the correlation coefficient for these data to be $+ 0 \cdot 88$ and the regression equation :—

$$\text{Number of thunderstorms} = 10 \cdot 4 + 0 \cdot 11 \text{ (Spot number)}.$$

Concerning the 27-day period Septer writes :—" Es sind schon Perioden 25 bis 27 Tage von D. O. Swiastskij wie auch die Periode $25 \cdot 8$ von V. Bezold und 27·5 von Ridder entdeckt worden, die sich sehr der synodischen Umlaufzeit der Sonne von 27 Tagen nähern ; man kann sagen, dass die Maxima der Gewitter auf der Erdkugel, die nach 27 Tagen folgen, durch irgend welche der Innern der Sonne gelegene Wirkungszentren hervorgerufen werden." There have been many investigations into the relationship between sunspots and rainfall, most of which have been inconclusive, but Clayton* summarizes them as showing that precipitation is greatest at or just after sunspot maxima at most tropical stations, in the North Atlantic and North Pacific, Southern Chile, the south coasts of Africa and Australia and at continental stations where summer rains predominate.

---

* 'Met Z.,' vol. 43, p. 229 (1926).

(D) *Nature of Ionization.*—In interpreting the wireless evidence it is well to remember that the effect of ionization in producing reflection, refraction, and absorption of the waves is measured, not by N the ionization content, but by N/m where m is the mass of one of the charged particles. We really therefore estimate $\left(\frac{N_e}{m_e} + \frac{N_\iota}{m_\iota}\right)$, where the e subscripts refer to electrons and the ι subscripts refer to ions. That N_e/m_e is greater than N_ι/m_ι has been demonstrated by measuring the polarization of the reflected waves. The influence of the earth's magnetic field is such as to make the ionosphere an aeolotropic medium and, owing to the difference in the group velocities of the two components, a single wireless pulse may be split into a doublet. Such an effect, which occurs both for Regions E and F, is only to be expected if the refractive process is effected mainly by electrons. The relation between the observed polarization and the direction of propagation relative to the earth's magnetic field demonstrate that the effective electrical charges are of negative sign.

(E) *Ionizing Agencies.*—Experiments on the occasion of the eclipse last year, in Canada, showed that the normal cause of ionospheric ionization is ultra-violet light from the sun. Now our picture of the structure of the ionosphere is that there are two main regions each of which is probably made up of two elements during the daytime. Probably the four components are associated with ionization potentials of different atmospheric constituents, atomic and molecular.

Possible abnormal ionizing agencies have been mentioned in (C). These are charged particles and their incidence, according to the measurements of Mr. Naismith and myself, is correlated with magnetic storms and with thunderstorms. They usually cause ionization in or just below the normal E Region, but the ionization gradient they produce is steeper than that produced about the same level by ultra-violet light.

---

* 'World Weather,' New York (1923).

(F) *Processes of Electron-Capture.*—During the night there is a steady decay of ionospheric electron content which, in the absence of abnormal influences, can be used to study the process of electron-capture. It is easier to make precise measurements for the upper region than for the lower region, but even for the former the evidence is conflicting. Eckersley, from a study of Marconi facsimile transmission, concluded that the usual recombination law for electrons and positive ions described the process, whereas I have recently found that such a law would not fit some of my measurements and that the law which describes the attachment of electrons to uncharged atoms gave a better agreement with experiment. Further work on this subject is necessary before a definite conclusion can be drawn.

Professor S. CHAPMAN, F.R.S. : I should like to express my admiration for the very clear exposition of the subject given by Professor Appleton I look upon the ozone and the ionization of the upper atmosphere as representing in a way an absorption spectrum of the sun's radiation. The sun's radiation clearly has different components which are absorbed at different levels, and forms a kind of spectrum in which the absorbing atmosphere is itself the spectroscope. Our problem is to interpret both what is the radiation of which that spectrum is formed, and what is the nature of the instrument which gives us the spectrum. The absorbing medium, the air, may for this purpose probably be regarded as composed solely of oxygen and nitrogen, but the various states of these two substances which are important in absorption may be somewhat numerous. Both gases doubtless occur in the atomic as well as the molecular state, and some of these four forms—atomic and molecular oxygen, and atomic and molecular nitrogen —probably have further varieties, with different and significant absorption. That is because the atoms and molecules have excited states, some of them being metastable ; the sun's radiation must continually produce excitation. Though the duration of the excited state for each individual particle may be brief, there will at any time be a certain number in each excited state. As far as concerns ionization, each of these different varieties is an independent constituent of the atmosphere, with its own absorption coefficient, and its own ionization potential, less than that for the normal particle of the same kind.

While the solar radiation of frequency high enough to ionize a normal atom or molecule is probably absorbed high in the atmosphere, there is a band of solar radiation of longer wave-length which can either excite the atoms and molecules, or can ionize those that are already excited. This second band has, in the aggregate, much more energy than the first, though the individual photons have less. The lower part of the ionized region may be ionized by this double process, though the actual absorption coefficients involved, which determine the level of each kind of absorption, are not yet known, either from theory or from laboratory experiments.

One point of interest may be mentioned, namely, that the ion content due to single-process ionization—that is, by the very short wave-lengths—must be in proportion to I, the intensity of the ionizing radiation, while if there is double-process ionization its intensity must be proportional to the product of the two ionizing radiations concerned, I_1 and I_2. If, therefore, the sun's radiation varies in the course of the sunspot cycle, as it is clearly shown to do by the magnetic evidence, and as Professor Appleton has shown to be probable from the radio evidence as well, it is not yet possible to infer from the variation of the ion content what is the variation of the solar radiation unless we know whether the ionization involved in the layer concerned is a single-process or a double-process ionization.

Among the particles which can be ionized more easily than the atoms and molecules in the normal states, I might also mention the negative oxygen ions formed by attachment of electrons to oxygen particles. The ionization potential of these negative ions is probably five or six volts, and so they can be ionized fairly readily, and form an additional source of free electrons. Negative-oxygen ionization may occur either by absorption of solar radiation or by an indirect supply of energy from the same source by collision with excited particles. The light of the night sky shows that excited particles exist throughout the night and are probably continuously formed, and there is a possibility that these may retard the decay in the electron content during the night.

The nature of this decay of electron content is a question of great interest. Do the electrons mainly recombine with positive ions, or do they attach themselves mainly to neutral particles? A question closely allied to that is the ratio of the number of electrons present to the number of ions; if electron attachment is important, then there must be many more ions than electrons. Our information on these points will, I think, come partly from the curves of daily variation of ion content at the various levels, and partly from the seasonal variations. I am making a detailed theoretical study of these points, and find that the situation as regards the daily variation is rather complex; but the seasonal variation of ions gives a clearer indication. If recombination is paramount, then the electron content in summer (n_s) as compared with that in winter (n_w) is given by the formula

$$\frac{n_s}{n_w} = \sqrt{\frac{I_s}{I_w}} = \sqrt{\frac{T_w \sin(\theta + \delta)}{T_s \sin(\theta - \delta)}},$$

where θ denotes the colatitude, δ the sun's maximum declination (23°), and T_s, T_w are the summer and winter values of the absolute temperatures of the atmosphere in the ionized layer. In our latitude this gives

$$\frac{n_s}{n_w} = 1\cdot84 \sqrt{\frac{T_w}{T_s}}.$$

The corresponding ratio when attachment is paramount is

$$\frac{n_s}{n_w} = \frac{T_w \sin(\theta + \delta)}{T_s \sin(\theta - \delta)},$$

or $3\cdot4\ (T_w/T_s)$ in our latitude.*

Professor Appleton has pointed out that difference in a previous paper, and his values, as indicated in his introduction, are about 2·2 for the E layer and 1·5 to 1·8, with less certainty, for the F layer. In both cases these values seem to indicate that recombination is more important than attachment. It may be added that if attachment is comparable in importance with recombination in the F layer, it should be paramount in the E layer, because of the greater density of oxygen particles to which the electrons could become attached. It may therefore prove that, when Professor Appleton's daily observations of ionization are continued, and when the average of a number of curves is taken, that he will find that this is the law of recombination in the F layer also.

Finally, I may remark upon the valuable eclipse observations obtained by Dr. Henderson† during the Canadian eclipse of August 31, 1932. These showed that n fell by approximately 60% during the optical totality, thus giving support to Appleton's view that this layer is ionized mainly by ultra-violet light and not, as I had suggested, by corpuscles. The possible existence of a corpuscular eclipse in some other layer is still, I think, worth further investigation. But with the assistance of Mr. J. C. P. Miller I have determined what the expected variation of n should have been, during the Canadian eclipse, assuming ultra-violet light as the sole ionizing agent, and assuming the degree of recombination (only) compatible with Appleton's daily-variation curves for the E layer.* The result agrees closely with Henderson's eclipse graph of *n*, and shows that there is no need to attribute any part of the residual 40% electron content at totality to another ionizing source. Similar calculations for the F layer† indicate that a much smaller eclipse effect was to be expected there‡ ; this was not examined during that eclipse, and remains to be tested by observations during later eclipses.

---

* The value of T_w/T_s is unknown, but seems unlikely to differ much (either way) from unity.

† 'Canadian J. Res.,' vol. 8, p. 1 (1933).

Professor C. T. R. Wilson, F.R.S.: I propose to say something about the possible connection between the ionosphere and thunderstorms. It is not perhaps generally realized how very great is the supply of energy from thunderclouds. The energy which is continually being spent in lightning flashes amounts, per square centimetre of the earth's surface, to two or three ergs per second. This is a thousand times as much as the energy required to produce all the cosmic radiation that enters the atmosphere, and would probably be sufficient to produce all the ionization of which we have evidence in the ionosphere.

I am going to assume that a thundercloud is of positive polarity. Such a cloud will pull electrons down from the upper atmosphere and, what is really of more importance, it may cause the field at lower levels to exceed the critical value required to make it effective in itself causing ionization. Professor Appleton has alluded to this effect in a recent paper. But in moderately high latitudes, where the effect of the magnetic field on vertically moving particles is negligible, the important critical field is not that corresponding to ordinary ionization by collision, but is only about one tenth as great. It is the smallest field that would be required to balance the energy loss of a beta-particle that is moving in the right direction and has the energy corresponding to the minimum rate of loss per cm. When this critical field is exceeded—it is a field corresponding at atmospheric pressure to 2000 or 3000 volts per cm.—corpuscular cosmic rays and their secondaries of limited range will have their range extended downwards, and increased ionization will result. If the potential of the upper part of the thundercloud exceeds about 10^9 volts, the critical condition for the acceleration of beta-particles will extend right down to the thundercloud. If the potential is increased beyond this point, a very large increase of current will correspond to a comparatively small increase of potential. The potential of the upper part of the thundercloud—the positive potential of the thundercloud—is thus not likely greatly to exceed 10^9 volts, and the conditions beyond this point approach instability; tend, that is, to give sudden equalization of potential between the upper atmosphere and the cloud.

---

* That is to say, the parameter σ_0 on which the daily variation depends (*cf.* 'Proc. Phys. Soc.,' vol, 43, p. 36, 1930), was taken as $\frac{1}{23}$.

† Taking σ_0 to be 1 or $\frac{1}{2}$.

‡ Slides were shown to illustrate eclipse effects on the E and F layers. The diagrams will be included in a future paper on the subject, now in preparation.

In equatorial regions the nearly horizontal magnetic field will prevent the initial acceleration of the secondary beta-particles by the electric field unless this electric field exceeds a certain limit—that required to balance the tendency of the magnetic field to turn the beta-particle laterally; this is a field of the order of 100 volts per centimetre. In consequence of this action of the magnetic field the effect of acceleration of electrons in helping ionization is in low latitudes confined to comparatively low levels, and the absence of this ionization in the upper levels adds something like a hundred million volts to the critical value of the potential at the top of the cloud.

The lower part of the cloud will only acquire big negative potentials when the positive potential of the upper part of the cloud has exceeded the critical value of the order of 10^9 volts; the lower part of the cloud may acquire a very big negative potential suddenly if the field above the cloud is suddenly destroyed.

Two important results would follow from such a destruction of the high positive potential at the top of the cloud. One would be a sudden raising of the negative potential at the bottom of the cloud, which might result in a discharge to earth. In addition, the sudden raising of potential at the bottom of the cloud enables runaway electrons, which have been produced within the cloud itself and have now no longer to traverse a retarding field above the cloud, to reach the upper atmosphere or to be bent round by the magnetic field and get to earth.

One point of interest to notice is that the additional hundred million volts or so, required where the magnetic field is mainly horizontal, may be one of the reasons why in equatorial and tropical thunderstorms lightning flashes are much less common between the bottom of the cloud and the ground, but are generally in the cloud. Again, the only evidence that exists so far that there is really penetrating radiation getting from the thundercloud to the ground has been furnished by Schonland's experiments in which he observes the coincidence between the kicks of a Geiger counter and atmospherics. His results suggest the production of fast beta-particles which are able to reach the earth under the action of the magnetic field, mainly at the moment of the sudden destruction of the electric field above the thundercloud.

The main production of what Eddington has called the runaway electrons must, of course, be within the thundercloud itself. There, as I have pointed out before, the beta-particles in the field, if this is at all strong, multiply enormously. Even taking what seems to be a minimum estimate, a single beta-particle starting upwards in the field within the cloud (which I am supposing to drive negative electrons upwards) would produce a thousand runaway electrons in passing through a distance corresponding to a potential fall of a hundred million volts. If the particle continued to move in the field through 10^9 volts, it would produce a quite impossible number—about 10^{30}—fast electrons, which would take more than the whole energy of the thundercloud.

It is quite possible that, within the thundercloud, as above it, the production of these fast electrons is the controlling agency which determines the limits of the potential difference attainable. It is worth noticing that in a very strong field the group of electrons produced from the single primary electron—an enormous number of electrons—would stick together as a localized group.

If the electrons are shot up from the thundercloud in large enough numbers, with the opposing field above the cloud removed, then we get the possibility of very important effects in the upper atmosphere. But what I have said about the very big effects produced by even a single original electron particle in the thundercloud suggests that we never can get a potential difference of even as much as 10^9 volts in a thundercloud. That limitation would prevent the runaway electrons from reaching to any great distance in the ionosphere, or after being bent by the magnetic field coming down as penetrating radiation. It is, however, conceivable that an electron produced within a thundercloud may acquire a bigger energy than corresponds to the electro-motive force of the thundercloud. For example, the particles at the tip of a progressing linear discharge—a discharge started by one of these groups of electrons—may be exposed to a force that is many times the primary electric field of the thundercloud. This force, if the electrons form the front of the advancing discharge, may continue to act on the leading electrons while they move through a distance comparable with the thickness of the thundercloud. In that way it is just conceivable that electrons may acquire energies very much greater than 10^9 volts, and that they may therefore give rise to very important ionizing effects in the ionosphere and also contribute to the penetrating radiation

Mr. T. L. Eckersley: Professor Appleton has clearly stated the three ways in which measurements can be made of the pulses reflected from the ionosphere.

I propose to discuss the latter two, that is, the measurements of the polarization characteristics and the intensity of the reflected pulses.

The apparatus used is shown on the slide. The principle employed is the following. The signals are received on two frame aerials at right angles to each other and are transferred by a rotating search coil (as with Bellini-Tosi direction-finding aerial) to the receiver and thence to the plates of a cathode ray oscillograph.

A circularly polarized incoming ray will induce equal e.m.f. in each of the two frames but these e.m.fs. will be 90° out of phase. If the two aerials are accurately tuned to the incoming wave, a rotating field will be produced in the goniometer. If, however, the aerials are mistuned, one, so that the current is 45° in advance of the e.m.f. and the other 45° in retard of the e.m.f., then when the e.m.fs. in the aerials are 90° out of phase, the currents in the two aerials will be in phase.

It follows that, for a circularly polarized wave, the search coil of the goniometer can be oriented until there is no e.m.f. induced. The zero position can be arranged so that it is $-45°$ for right-hand polarization and $+45°$ for left-hand polarization.

The plane polarized direct signal is of uniform strength when the direct ray angle bisects the planes of the two aerials.

By this method we have a means of distinguishing the polarization of the down-coming waves.

An auxiliary source of high frequency oscillation is used in conjunction with the receiver in order to facilitate the correct adjustment of the two aerials, and for the further purpose of measuring the intensity of the reflected pulses.

Observations—wholly visual—have been made on wave-lengths between 45 and 90 m. mainly in the daytime. The transmitter of 200 watts power was situated at Chelmsford and the receiver at Broomfield, 1·5 km. away.

One of the most striking features of the results is the extraordinary variability of what may be called the ionospheric weather. Each day brings a different set of conditions and a large number of systematic measurements are required to disentangle the main features.

On a great many occasions split pairs have been observed. This is in accordance with the magneto-ionic theory which predicts the splitting of a plane polarized emitted wave into two components of opposite circular polarizations if the electrical carriers are mainly electrons. The normal split pairs are definitely oppositely polarized. In general, the right-hand one is least delayed and the left-hand most delayed although occasions have been observed when the opposite occurs.

On rare occasions there are very definite triplets, with one right-hand and two left-hand following, all sufficiently close to ensure that they belong to the same system.

Triplets have been observed on five occasions during the last five months on wave-lengths between 55 and 60 metres.

A point worth mentioning is the frequency with which E and F reflections are recorded simultaneously.

This definitely suggests partial penetration and partial reflection, since, on some occasions, both right- and left-hand E reflections occur simultaneously with F reflections, so that this fact is definitely in contradiction with the ray theory which requires that a ray shall either completely penetrate or be completely reflected.

Another characteristic which may be stressed is the occasional spreading of the echoes, usually just before the echo in question is lost by penetration of the layer. This multiple splitting has been observed on many occasions and is all of one polarization ; it is, in general, not split and split again into alternate left and right components.

This effect of spreading is, I think, mainly due to dispersion as suggested by Pendl* and occurs when there is a rapid change of group and phase velocity with frequency.

This rapid change is specially great for frequencies close to the penetration frequency, which accounts for the fact that spreading is generally observed just before the ray in question penetrates the ionosphere.

Polarization characteristics vary very considerably from day to day.

The results of a large number of observations indicate, however, a certain regularity in their incidences.

Thus on 60 m. left- and right-hand polarizations are nearly equally prevalent, as the following figure of a series of observations during December and January and part of February this year, shows. On shorter wave-lengths the right-hand polarizations are predominant and on longer the left-hand ones.

Relative Attenuation of Left- and Right-hand Polarized Waves.

This is illustrated in the attenuation curves, which indicate that :—

(1) On 60 m. in daytime the R and L rays are nearly equally attenuated.

---

* Elect. Nachr. Tech., vol. 10, 1933, pp. 75–94.

(2) There is a progressive decrease in attenuation as we approach sunset where the reflection coefficient approaches 0·5.

(3) An almost stationary value of attenuation coefficient of right-hand component over a range of about 4 to 6 megacycles.

(4) The absence of left-hand components on frequencies over 5 megacycles, either on account of electron limitation or excessive attenuation of the left-hand ray.

From these results, we can obtain a measure of the daytime reflection coefficient on these waves. Thus, from the known characteristics of the transmitting aerial and the transmitting current, the emitted wave, if it spreads out uniformly and is reflected at 250 km. height without loss should give a field intensity of about 350 μv/m. It will be observed that the measured fields converge to the neighbourhood of about half this value after sunset when the reflection coefficient should be nearly 1/2. The reflection coefficient is of the order of 1/60 for mid-day, winter reception on 60 m.

The loss of energy on reflection is largely, if not completely, determined by the collisions of the electrons or positive ions set in motion by the wave, with the neutral particles in the atmosphere. (Apart from partial reflection and transmission which is confined to a very narrow band of frequency.) The evidence then implies that there is quite considerable energy loss from this cause in the F layer, and that this loss is greatest for the left-hand component. The reasoning is as follows :—

Both the left- and right-hand rays pass through the E layer to the F, are reflected there and pass a second time through the E layer to earth again. Both rays have to pass through the whole of the E layer twice. For each element of path in the E layer the specific attenuation of the right-hand ray (assuming that the electrons are predominant) is some three times (for λ 60 m.) that of the left-hand ray. Since the two rays follow the same path the total attenuation of the rays in the E layer should be in the ratio of 3 to 1 in favour of the left-hand component. The fact that, however, attenuation of the left-hand component is equal to that of the right implies that the reflection of the left-hand component at the F layer is much more feeble than that of the right-hand component.

This decrease in reflective power for the left-hand component might possibly be due to electron limitation, but partial penetration associated with electron limitation is confined to a very narrow band of frequencies and this phenomenon of poor reflection of the left-hand component is spread over quite a wide band of frequencies.

In these conditions the total attenuation coefficient of the left-hand ray in the F layer is at least two thirds of the total measured attenuation, *i.e.*, approximately 36 to 40 decibels at mid-day. It follows that the F layer attenuation of the left-hand ray is at least 24 to 27 decibels, *i.e.*, the reflection coefficient of the left-hand ray at the F layer is as little as 1/16 to 1/22 (for vertical transmission) at mid-day.

Again, if the loss of reflection in the F layer were due to partial penetration we should expect the F layer attenuation to get greater towards sunset as the density decreases, and the penetration frequency for the left-hand ray is more nearly approached, the opposite is, however, the fact.

The conclusion that the left-hand ray is largely attenuated in the layer may be avoided if we assume that the effective electric carriers which perform the function of modifying the refractive index of the lower E layer are positive ions, single ionized oxygen molecules, for instance.

In this case there would be inappreciable differential absorption between the left- and right-hand components in the E layer (since the critical rotation frequency for these carriers is very small) and the whole of the attenuation on 60 m., where the right- and left-hand components are nearly equal, might be attributed to E layer absorption. If this were so the right- and left-hand attenuation should be equal on all wave-lengths 45 to 90 m., and observation shows that this is by no means so, the left-hand component being predominant on the waves longer than 60 m. frequencies (5 kilocycles) and the right-hand component being predominant on the shorter wave-length. We should thus again have to attribute the difference in reflection of the left- and right-hand components to differential absorption in the F layer.

That the left-hand component can be much more absorbed in the F layer than the right must be explained. Although the specific attenuation of the right-hand ray per unit length is greater than that of the left-hand component the deeper penetration and longer path of travel may easily compensate for this effect; indeed, in general, the attenuation, although not exactly proportional to the group time of travel of the pulse, varies more or less in proportion, and the undoubted fact of the longer group time of the left-hand ray definitely encourages the supposition or possibility of a greater absorption of this ray.

These attenuation results are rather difficult to explain on any preconceived motion of the air density in regions above 100 km.

Thus, if we assume that the dissociation process for O_2 is nearly complete at 120 km. height so that the oxygen atom is predominant above this height and that the height of the homogeneous atmosphere for this case is 23 km., it is possible to calculate the collision frequency above this point. Such a distribution would then be in agreement with Chapman's* estimate of the density at 200 km. height.

With any reasonable distribution of ionic density the attenuation in this region should be practically nil. To account for the attenuation, a 50- to 100-fold increase in collision frequency would be required in these regions such as might occur if dissociation were complete at 100 km. and H above this region were 20. The collision frequency would then vary from about $3 \cdot 5 \times 10^5$ at 100 km. down to about 2×10^3 at 200 km.

There is considerably more evidence of a more qualitative sort, from the results of long distance transmissions, which supports the view that there is appreciable F layer attenuation. This evidence is discussed elsewhere.†

Layer Densities.

The subject of the maximum layer densities and their determination by observing the critical penetration frequencies has been dealt with at length by the other speakers.

The results of individual measurements are very variable. A large amount of statistical material is required to show the general run of the diurnal and seasonal variation.

I have attempted to obtain a great deal of material by a quick method, which may perhaps be lacking in accuracy, but which rapidly gives a large number of results which can be analysed statistically. The method is based on observations of local sending stations and a determination by direction finding methods of whether the signals from the given station skip over the receiver or not.

The skip phenomena is well known. For a given range between sender and receiver and an assumed layer height it is possible to calculate the electronic density which is just sufficient to return the ray to earth at the receiver. If, at the moment of observation, the station is within the skip distance, *i.e.*, receives little energy reflected from the layer, the density must be less than this critical amount; if beyond the skip, the density is greater than this amount.

The results are plotted on the slide. The arrows represent individual observation: those pointing upwards represent the fact that the density is greater than this value at the time of observation and *vice versa* with the arrows pointing downwards.

---

* 'Proc. Roy. Soc.,' A, vol. 32, p. 353 (1931).

† 'J. Inst. Elect Eng. London,' vol. 71, p. 405 (1932).

The upper limit of the arrows pointing upwards and the lower limit of the arrows pointing downwards should define a single curve giving the electronic density time relation (*i.e.*, the diurnal variation of the density).

Actually on account of the variability of the observations the arrows interpenetrate a little, but a curve can be drawn such that it lies above about 80% of the A arrows pointing upwards and below about 80% of the B arrows pointing downwards.

The method is based on perhaps rather doubtful assumptions, but it is satisfactory to note that the results obtained are checked by the more direct method of the penetration frequencies for normal incidence.

The results are shown for the Winter, Spring, Equinox, and Summer of 1932 and show the seasonal as well as diurnal variation of the density.

The ratio of maximum summer mid-day density to maximum winter mid-day density observed is 1·9 to 1.

The curves actually drawn are theoretical, *i.e.*, give the calculated maximum density, assuming, after Chapman, that the ionization is produced by a monochromatic ionizing radiation travelling in straight lines from the sun and that there is a constant rate of recombination.

The curves are drawn for $\sigma = 0\cdot5$ where σ is Chapman's constant.

$$\sigma = \frac{1}{1\cdot37 \times 10^4 \alpha N_0},$$

where α is the recombination coefficient and N_0 is the maximum electronic density which would be reached if the ionizing agent remained on indefinitely at vertical incidence.*

The average recombination coefficient, assuming recombination is the predominant factor, can also be determined from the times of disappearance of the right- and left-hand rays in the experiments and observations made during January and February this year.

The values obtained lie between $1\cdot3$ and $1\cdot5 \times 10^{-10}$ and do not differ greatly from those already given.

---

* The values for the best fit are

$$N_0 = 1\cdot61 \times 10^6$$
$$\alpha = 0\cdot9 \times 10^{-10}.$$

Measures of the recombination coefficient by the facsimile method described in Eckersley (*loc. cit.*), are $0\cdot93, 0\cdot95, 0\cdot67 \times 10^{-10}$ and agree, except the last, with the value determined here.

Undoubtedly the density does not always drop off smoothly at night; many occasions have been observed when there has been a sudden increase, but I am inclined to the view of a regular decrease at night after the ultra-violet radiation has ceased together with occasional increases due to solar emission of charged particles. The average recombination coefficient determined as above is likely to be a little too low.

The theoretical densities according to Chapman's theory have been calculated by a numerical method of Millington and the following contours of equal density are shown in the slide; they are calculated with $\sigma = 0 \cdot 5$ for Summer, Equinox, and Winter. The validity of this chart has been tested by many observations and measurements of long distance short-wave transmissions.

These results have some bearing on the question of the mechanism by which the electron density decreases during the night when the ionizing agency is removed. Two alternatives have been proposed.

One, that the number of electrons decreases in virtue of their recombination with positive ions and the other in virtue of the attachment of the electrons to neutral ions. Chapman has shown that the ratio of the maximum summer to maximum winter ionic density according to the former theory should be 1·84 for 50° latitude and according to the latter theory $(1 \cdot 84)^2 = 3 \cdot 4$. The observations agree with the former theory, and the further agreement of the ionic density charts with the interpretation of long distance short-wave transmission results, although of much less weight, tend to confirm the supposition that recombination is the major factor determining the rate of decrease of electronic density at night.

Mr. R. A. Watson Watt: I hope to be brief, because I should be most disappointed if time permitted only those whom I might call the professional ionospherists to take part in this discussion. Professor Appleton has dealt particularly with the fine structure of the ionosphere in vertical section; I should like to say a very few words about the structure in horizontal section. Professor Appleton has dealt with the striped structure; I should like to deal with the spottiness.

As a contribution to the examination of thunderstorm effects on the ionosphere, my colleagues and I at Slough—and here I think I should, if a little invidiously, mention especially Mr. Lutkin—have compared the records of ionospheric conditions with detailed observations of thunderstorms. Professor Appleton and Mr. Naismith have already reported the results of a statistical investigation made at Slough, but I think it is of interest to quote one or two particular examples in which local thunderstorms were studied in relation to the instantaneous state of the local ionosphere.

The first very definite observation was made by Mr. Lutkin in October, 1932, and was of this nature A thunderstorm with locally audible thunder was observed in association with a condition in which, after an exploring frequency had been reached at which no reflection was observed from the upper region, reflection at a still higher frequency reappeared from the lower region. That is, there was a temporary state in which ionic density in the lower E region ran up to something of the order of 8×10^5 electrons per cubic centimetre, while the density in F region was still down in the neighbourhood of 3 or 4×10^5. The statistical work reported by Appleton and Naismith followed very soon after this observation, while a further example of relatively short-duration variation in ionic content was found in April of the present year.

A continuous record during the afternoon of reflections of a 6 megacycles/sec. pulse showed several periods, of duration between 3–5 secs. only, when waves of this frequency were being reflected from the E region. The critical frequency at noon was 3·1 megacycles/sec. and assuming that this indicated the normal peak ionization for that day these very short duration increases represented momentary increases of at least threefold in density. These increases were of such short duration that there was no time to manipulate the ordinary hand-operated camera, which was ready for operation.

This phenomenon is not at all infrequent under summer conditions. As illustrating other numerical measures of the rate of increase of ionization in these abnormal and probably thunderstorm conditions, one might cite a three to one increase which took place in a single hour in a May observation ; a two to one increase within half an hour, on an occasion when thunderstorms were definitely observed within 25 kilometres ; and a series of observations which cover the work of the present month of June. The full curve (shown on the slide) represents the number of atmospherics per hour recorded on a particular instrument of fixed sensitivity This curve is believed, from our general experience of relative levels of disturbance, to give a fair indication of the thunderstorm activity within 50 kilometres of the recording station. The individual dots indicate the electron content of the lower layer at noon in the daily routine observations at Slough. The shaded areas or individual vertical lines represent the comparatively brief occasions during which the electron content of that layer rose to abnormal values — values of more than $8 \cdot 5 \times 10^5$ electrons per cubic centimetre for one height, or 12×10^5 for the other height of ordinate, these two values corresponding to the two exploring frequencies used in our recording work. I think it is clearly established by this series of observations that the sudden appearance of these bursts of abnormal ionization just precedes the arrival of thunderstorms in the area within 50 km. of the recording station.

The method of removal of these abnormal densities from the zenith of the receiving station is something which calls for very close examination. Professor Wilson's significant remark that the secondaries belonging to a single runaway electron maintain themselves in a homogeneous group suggests that these may be the actual ionic cloudlets produced by such a mechanism, and observation over a chain of stations not very far apart might establish the method of dissipation or drift of these extreme concentrations.

There is implicit in these ideas of thunderstorm effects on the ionosphere, a suggestion for a very much closer examination than has come to my knowledge of the relations between magnetic disturbance and thunderstorm activity. It is clear that we should expect the abnormal conductivities in these abnormal bursts, produced by identifiable terrestrial sources, to be reflected in the terrestrial magnetic observations.

I feel that there is risk of an *embarras de richesse* in our available explanations for the ionization content of the ionosphere We appear to have in ultra-violet light alone a sufficient source for all the electrons in the ionosphere, and in tropical and other thunderstorms an equally sufficient explanation for the whole content. We may find it difficult not to explain too much between these two sources.

Arising out of Professor Appleton's introductory address, I should like to refer to one quite inconclusive examination which has just been made of a possible 27-day recurrence tendency in thunderstorms. Here (fig. 2, shown on screen) is an application of Bartel's method of plotting, vertically below one another, 27-day epochs of particular types of data. We have plotted the number of atmospherics per minute received at Slough in particular and defined circumstances. In the lower one we have plotted the intensity of these atmospherics. The data are, as I have said, inconclusive, because neither the integrated intensity nor even the number of atmospherics crossing a given threshold is a unique indication of the thunderstorm activity within a given radius. This is due to the variation in the intensity of the received atmospherics with the variation of effective reflection coefficient, which is in turn due to that very variation of ionic content aloft which we are discussing. But this possible secondary effect of a possible independent 27-day turn in ionic content does not, I think, invalidate the suggestion that fig. 2, based on records from a comparatively insensitive instrument, an instrument, that is, which probably neglects the remoter sources more strongly affected by varying reflection coefficient, sufficiently supports the proposal for further study of the 27-day recurrence tendency in thunderstorm activity. That further study would require closer investigation of the thunderstorm distribution than I believe to be economically practicable by any method other than that depending on the radiotelegraphic location of sources of atmospherics; this method must, however, first be further developed, in relation to the interpretation of the observed intensities of atmospherics, than has yet been found possible.

Turning to the questions of fine structure in the vertical already discussed by Professor Appleton, I should like to remark that I believe that there is more than one sharply marked level of maximum density in the E region itself. In many of the cases of brief temporary increase of density which I have already outlined, we find that the increases take place—quite frequently in turn—at two or three well-defined equivalent levels each differing by some 10 or 15 km. from its neighbour. Thus, two days ago, in the period 0800—1100 G.M.T. of June 20, 1933, the equivalent height of sustained reflection for exploring pulses of 4 megacycles/sec. was 110 km. But the sudden bursts of increased density which I have discussed took place, in (*a*) 10 first-order reflections with none of second-order at 97 km., in (*b*) 4 of first-order accompanied by 7 of second-order, at 112 km. and (*c*) in 7 of first-order accompanied by 1 second-order reflection at 122 km. The second-order reflections gave values of 117 and 127 km. respectively. The maximum departure from any of these means was 3 km., which represents the order of accuracy of measurement. While it would be possible to devise a partial explanation in terms of magneto-ionic splitting, the circumstances do not favour such an explanation, which would demand a change from complete absorption of one component to complete absorption of the other, and *vice versa*, at repeated intervals of a few minutes. No installation sufficiently elaborate to test this improbable explanation has yet been set up, although there would be no difficulty in meeting a demand for it. Meanwhile, I favour the view that such data reveal additional fine structure in the vertical section.

Mr. J. A. **Ratcliffe**: During the last twelve months automatic records of wireless signals returned from the ionosphere have been made at Cambridge by means of an apparatus designed by Mr. E. L. C. White. The apparatus is constructed from components lent to us by the Radio Research Board. In the records the effective height is plotted against the time of day, the wave-length remaining constant (P′*t* curves), and provision is made for the automatic registration of the polarization of the downcoming waves. A preliminary survey of these records reveals several points of interest for this discussion.

The most striking feature is the phenomenon, noticed previously by several workers, of a nocturnal increase in the E region ionization. This generally occurs without any corresponding increase in the F ionization, and has even been noticed on occasions when the F ionization is abnormally weak. There is a close correlation between the occurrence of the nocturnal E region and disturbed magnetic conditions. Anomalous magnetic behaviour is also occasionally associated with unusually low ionization in the F region. The presence of intense ionization in the E region for an unusually long time after sunset is correlated with the occurrence of thunderstorms.

The records show abundant evidence of the intermediate region. In the winter months this region is often more intensely ionized, in the day, than is the E region. Thus, in November and March echoes from the intermediate region appear ($\lambda = 150$ m.) for about an hour or two near sunrise, before the E echo is established and again near sunset between the disappearance of the E echo and the appearance of the F echo. In December, January, and February it is quite usual for the intermediate region ionization to remain more intense than that of the E region throughout the day, so that (on 150 m.) the echo is returned all day from the intermediate region, with occasional spasmodic reflections from the E region.

The effective heights recorded for the E region are remarkably constant, they do not deviate from the mean (105 km. for a wave-length of 150 m.) by more than five per cent. which is only just greater than the limit of accuracy of measurement of the records. The effective heights recorded for the intermediate region are more varied, falling between 120 and 180 km. at different times of day (150 m.).

The nocturnal E region always has an effective height equal to that of the daytime E region (105 km.) within the order of accuracy mentioned above. The result is that in the winter months it is quite common to have reflection from the E level (105 km.) at intervals during the night, but from the higher intermediate region (approx. 140 km.) during the day.

On one occasion in December, 1932, when the usual behaviour was for the intermediate region echo to be present during the day, it was found that an F region echo alone was present throughout the day, thus indicating a shortage of ionization in the intermediate region. On the previous night there had also been a marked shortage of F ionization; it thus appears probable that the same source of ionization produces both the F region and the intermediate region. On this occasion the nocturnal E echo was very pronounced, so that it seems reasonable to suppose that the ionization of the E region is due to some different cause.

The nocturnal and daytime E regions observed during the winter months at a height of 105 km. are presumably due to similar causes, the very much greater absorption of the daytime echo may be due to a downward extension of the ionization of the intermediate region to levels below 105 km.

Polarization measurements indicate that free electrons are effective in the F, the intermediate, and the E regions, as also in the absorbing region, down, at least, to a level where the frequency of collision between electrons and molecules is about 5×10^6 per second.

The diurnal variation of the ionization density of the F region appears to be complicated. Rukop and Paul found an " evening concentration " of ionization about 2000–2300 hours last August, we have confirmed this, and find that it occurs about 1900–2200 hours in May of this year. There is also reason to believe that there is an increase of ionization near midnight, which, on a critical wave-length, often causes the echo to return during the midnight hours. Our records do not support the hypothesis of Elias, who considers that this returning echo is due to the presence of a still higher layer (his Layer II.).

There are signs that the ionization of the E region is less than it was a year ago, but there is no indication of a similar decrease in the ionization of the F region. This adds to the previous evidence that the E and F region ionizations are due to different causes.

Professor F. A. LINDEMANN, F.R.S.: Throughout the discussion we have been dealing with equivalent heights, defined as the heights calculated on the assumption that the group travels with the velocity of light. In order to correlate our information about the ionosphere with our geophysical knowledge, it is essential to know the real heights at which reflection of the wireless waves take place, for all our observations and calculations of temperature, density, constitution and so on, of the upper air, refer to real heights. It is in this connection that it is perhaps worth while to emphasize that there does not seem to be any conclusive evidence that the equivalent heights are closely related with the actual heights.

It is well known that so-called wireless echoes of groups of signals have been observed tens of seconds and sometimes even minutes after they have been sent out. So far, only two explanations of this phenomenon have been put forward. The first is that the waves are reflected by a cloud of charged particles coming from the sun bent into an appropriate form by the earth's magnetic field. This explanation does not seen to hold water. In the first place there would have to be some 10^5 to 10^6 particles per cubic centimetre. If they were of equal sign their mutual repulsion would be so excessive as to render the formation of the cloud impossible. If there were equal numbers of opposite sign then the calculations which justify the assumption that such clouds are likely to exist, would not be correct. But whether this is so or not, this explanation may be ruled out for the following reason. If the wireless signals really had to travel distances measured in millions of kilometres corresponding to the delay before the echo returns, they would be so weak as to be imperceptible. This objection is usually met by maintaining that the cloud of particles assumes such a form as to focus the wireless waves back upon the earth, thus overcoming the effect of the inverse square law. In reply to this it must be pointed out that such a focussing action can only occur if the curved surface of the cloud is defined with an accuracy comparable with the wave-length of the signals in question. It will scarcely be maintained that these particles coming from the sun with all sorts of velocities are bent into a curve with a radius of millions of kilometres which is defined to an accuracy of a few metres. A further phenomenon is that the period of delay before the echo returns sometimes varies within a few minutes by a considerable number of seconds. If one wishes to maintain the cloud theory one must be prepared to assert that the cloud surface moves millions of kilometres within a very few minutes, all the time of course maintaining its sharply defined surface accurate to a few metres These objections are so grave that the electron cloud explanation of the echoes may be rejected.

Unless we are content to believe that the echoes are produced by individuals signalling to us from the moon we are therefore driven back to the other explanation, namely, that the group velocity is sometimes abnormally retarded in the ionosphere by special distributions of ionic density which produce anomalies in the dispersion sufficient to delay the return of wave groups by times measured in seconds. This theory also has its difficulties but at present it seems to be the only one we have. Whatever the truth there is no doubt that retardations can and do occur. Hence it would appear reasonable to be on our guard lest similar, but infinitely smaller, anomalies in the dispersion be the cause of the retardations, of the order of thousandths of seconds, which we observe in the ordinary measurements of the height of the various ionized layers. If this were so, it might well be, for instance, that the two main layers instead of being separated by a hundred kilometres or so, as is generally assumed *sub silentio*, are really quite close together and merely represent more or less typical changes in the ionic density gradient. Questions such as these must be carefully considered and determined before we expend too much time or effort in endeavouring to relate the existence of these layers with other physical phenomena.

It is perfectly true of course, that the height of these layers is always referred to as the equivalent height and that nobody has ever claimed that they actually represent real heights. In discussions such as these, however, this fact is apt to be overlooked and one tends subconsciously to identify equivalent heights with real heights. It may be that some of the difficulties and apparent contradictions which one meets with in endeavouring to elucidate these phenomena would be reduced and perhaps eliminated if this factor were borne in mind.

The Transmutation of Lithium by Protons and by Ions of the Heavy Isotope of Hydrogen.

By M. L. E. Oliphant, Ph.D. (Messel Research Fellow of the Royal Society), B. B. Kinsey, B.A. (Trinity College), and Lord Rutherford, O.M., F.R.S. (Cavendish Laboratory, Cambridge).

(Received August 1, 1933.)

In a previous paper* on the transmutation of the elements by protons we reported some measurements of the disintegration function for thin films and for thick films of the element lithium, and gave the results of observations of the effects produced by molecular ions. In this communication we shall deal with the ranges of the particles emitted from lithium when bombarded by protons, and we shall also report the results obtained when that element is bombarded by ions of the heavy isotope of hydrogen.

The apparatus was described fully in the previous paper, and with the exception of changes in the windows which allow the particles produced to enter the counting chamber remains as before.

---

* 'Proc. Roy. Soc.,' A, vol. 141, p. 259 (1933).

Proton Disintegration.

Cockcroft and Walton* reported that the bombardment of lithium by a mixed beam of hydrogen ions gave rise to disintegration particles of two ranges. At about 400 kv., particles were observed with a range of about 8·4 cm. in air, while when the absorption was reduced to the equivalent of about 2 cm. of air† a second range began to appear. At about 6 mm. total absorption the number of particles counted had risen to twice the number in the longer range group. The 2 cm. group showed a gradual rise and did not appear to represent a very definite range. The numbers were still rising at the smallest absorptions used. Kirchner‡ reported that he could find no trace of this second group in Wilson chamber photographs, though the 8·4 cm. group and also the particles resulting from the disintegration of boron were photographed. In view of these contradictory results, and in order to obtain more accurate data for purposes of discussion, the ranges of the particles from lithium were re-examined in some detail. For this purpose windows of mica of stopping power 0·948 cm., 0·575 cm., and 0·15 cm. were used. This last window showed vivid interference colours and was mounted by means of a solution of celluloid over ten holes 0·35 mm. in diameter grouped together in a circle of about 3 mm. diameter. The beam of particles entering the counter was canalized by means of diaphragms so that the divergence was less than 15°. Fig. 1 shows the absorption curve obtained with these windows. The long range group is found to possess a mean range of 8·4 ± 0·2 cm., in agreement with the results of Cockcroft and Walton, while in addition there is present a very definite group of mean range about 1·15 cm., and probably also a third group of mean range about 6·5 mm. It is difficult to be quite certain of the ranges of these two short groups, for there is no source of radioactive α-particles available with which they can be compared as can the longer range group, which was checked against the 8·6 cm particles from Th C. However, we believe the error cannot be greater than about 2 mm. The numbers of particles in the three groups are approximately in the ratio 1 : 1 : 0·5, for the 8·4, 1·2 and 0·7 cm. groups respectively. In the paper which follows this Walton and Dee, p. 736, have given evidence from Wilson chamber photographs which supports these conclusions.

Confirmatory evidence of the presence of two short ranges was obtained by observation of the sizes of the deflections produced on the oscillograph by particles entering the counting chamber through varying thicknesses of absorbing material. Fig. 2 is the graph obtained by plotting the maximum size of deflections produced against the absorption, for a chamber 1·5 mm. deep. It is seen that the curve exhibits two maxima at about the expected absorptions. The maximum size of the deflection was identical with the maximum deflection produced by the α-particles from polonium, showing that the particles concerned carry two charges.

---

* 'Proc. Roy. Soc.,' A, vol. 137, p. 229 (1932).

† 'Nature,' vol. 131, p. 23 (1933); 'Proc. Roy. Soc.,' A, vol. 137, p. 229 (1932).

‡ 'Naturwiss.,' vol. 21, p. 32 (1933).

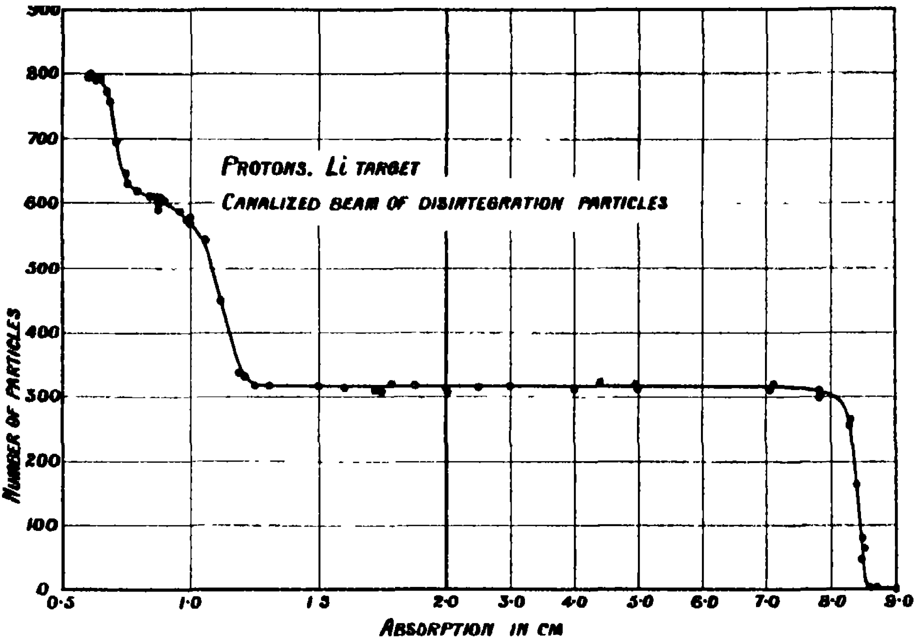


Fig. 1.

Disintegration by Molecular Ions.

The ranges of the particles ejected from lithium by bombardment with molecular ions of hydrogen were found to be identical with those produced by protons, and in fact the molecular ion is equivalent exactly to a pair of protons each with one half the energy corresponding to the accelerating voltage. However, there are present always a few particles per minute with ranges beyond the limit of 8·6 cm. found for the protons. For instance with 160 kv. accelerating and a molecular beam of about 3 μa hitting the target there are present about 4 particles per minute the range of which is greater than 10 cm. We shall refer to these particles later.

Disintegration by Ions of the Isotope of Hydrogen of Mass 2.

Lawrence, Livingston and Lewis* have reported that at energies of 3 × 10^5 to 1·2 × 10^6 volts ions of the heavy isotope of hydrogen are able to disintegrate many elements, and for lithium they found them to be about 10 times as efficient in producing disintegration particles as were protons. The beam which they use is one of (H^1 — H^2) molecular ions, so that it consists of equal quantities of protons and of H^2 nuclei. They find that there are present in the products of the transmutation of lithium, particles of 8·2 cm. range and others with the remarkably long range of 14·5 cm. The first group they ascribe to the protons, identifying it with the 8·4 cm group described above.

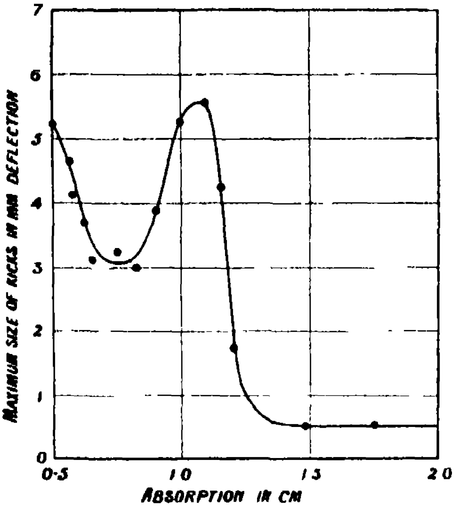


FIG. 2.

Professor G. N. Lewis has very kindly presented to us a sample of heavy water which he estimates to contain 93% of its hydrogen as H^2 We have prepared hydrogen gas from a portion of the sample, and about 20 c c. of this gas was mixed with 2 litres of purified helium. This was done, firstly, in order to conserve the H^2, as we found that the number of particles released from lithium at 150 kv. was so large that all measurements could be made with 1% of the activity. Secondly, the helium, by reason of its high excitation and ionization potentials, produced complete dissociation and strong ionization of the hydrogen in the discharge. The result of this is that magnetic bending of the accelerated beam is able to effect a complete separation of the $(H^2)^+$ ions from ions of ordinary hydrogen, and the effect of the heavy isotope as a disintegrating agent can be studied in the absence of any extraneous effects. The beam of ions obtained in this way was usually of the order of 0·5 μa. Accidental inclusion of ordinary hydrogen, partly from the walls of the discharge tube, gave rise to a proton beam of about 2 μa which could be brought on to the target by alteration of the magnetic field in order to compare the effects produced by the heavy isotope with those due to protons.

---

* 'Phys. Rev.,' vol. 44, p. 55 (1933).

The Disintegration Function.

Fig. 3 shows the variation in the number of particles ejected from Li with change in the accelerating voltage. The two full curves represent the results for two different absorptions between the target and the counting chamber.

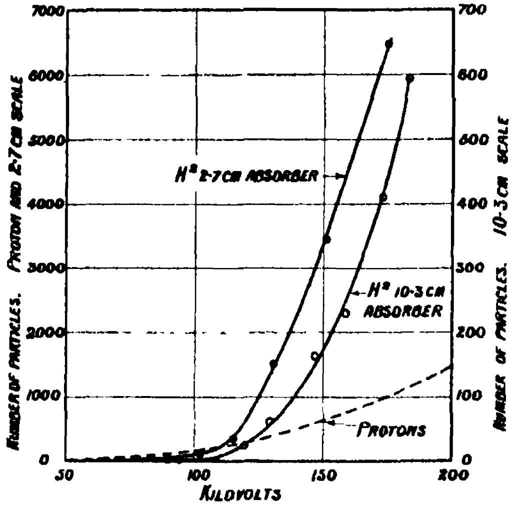


Fig. 3.

It is seen that the curves show approximately the same sort of increase with voltage. The dotted curve in the same graph shows the corresponding variation for the same proton current, and it is evident that while the heavy isotope does not produce an appreciable effect till a considerably higher voltage is reached, the subsequent rate of increase is so much greater that at 140 kv. the number of particles passing through the 2·7 cm. absorber is more than four times the number found with protons.

The Range of the Particles.

The range distribution among the particles produced from Li by $(H^2)^+$ ions is shown in fig. 4. It is seen that there is a very definite group of particles with an extrapolated range of 13·2 ± 0·2 cm. The mean range of this group is probably about 13 cm., corresponding to an energy of $11 \cdot 5 \times 10^6$ ε volts. There is also present a second group of particles possessing ranges distributed almost uniformly between zero and about 7·8 ± 0·2 cm., *i.e.*, with energies up to $8 \cdot 3 \times 10^6$ ε volts. The dotted curve in the same figure is the end of the corresponding range curve, obtained during the same experiment and under identical conditions, for the particles produced by the proton beam. The difference in range is clearly shown, and it is evident that unlike the 8·4 cm. particles produced by protons the shorter range group produced by H^2 ions is complex. The 13·2 cm. group contributes about 0·1 of the total number of particles counted at the smallest absorption we have used. It will be seen that these observations are completely confirmed by Walton and Dee, p. 739, by photographing the tracks of the particles in a Wilson chamber.

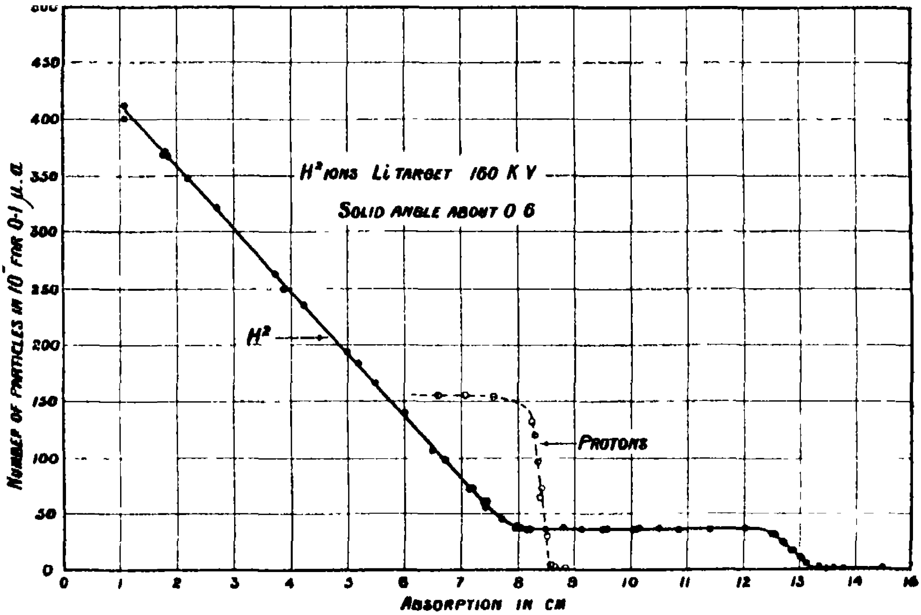


Fig. 4.

The range measurements were made by counting the kicks produced by the particles with a thyratron "scale-of-two" counter* and by observation of the oscillograph records obtained photographically. The photographic method gives a valuable check on the automatic counter, and in particular allows the entry of a new range to be seen at once by the presence of large kicks, quite apart from any number count. The maximum size of the deflections observed for both groups of particles was the same as for the particles from polonium, so that it is probable that they both consist of α-particles.

---

* Wynn-Williams, 'Proc. Roy. Soc.,' A, vol. 136, p. 312 (1932).

A check on the linear character of the range-number curve for the shorter range group is obtained by observation of the fact that the oscillograph records show the same number of large kicks at all points along the linear slope of the curve. This indicates that the same number of particles end in the chamber whatever the absorption, *i.e.*, that the slope of the curve is constant. It should be pointed out that we are using a solid angle of collection of about 0·5 in these experiments, and this will distort the range-curve slightly, making it fall off more rapidly at first than it should for a parallel beam of particles. This would mean that the true curve is probably not linear as we find it, but slightly convex to the axes. Such an error will be small and will not seriously affect our later conclusions. A careful search for particles of range greater than 13·2 cm. was made at accelerating potentials up to 200 kv. with a chamber 1 cm. deep so as to pick up any protons which might be present, but no particles could be detected in quantity greater than about 1/500 of the α-particles produced, which was our natural effect for the chamber. The fast particles resembling protons in their behaviour which were found by Walton and Dee, p. 739, probably appear only at higher bombarding energies.

Determination of the Concentration of H^2 *in a Sample of Hydrogen.*

It is now obvious that the few particles with range greater than 10 cm. detected when a molecular beam from ordinary hydrogen strikes a lithium target are probably 13·2 cm. particles produced by $(H^2)^+$ ions. Various estimates have been made of the concentration of H^2 in hydrogen, but it is probably of the order of magnitude of 1 part in 10,000. The production of the long-range particles by bombardment of Li is thus a very sensitive test for the presence of H^2 and it is possible to follow its concentration by electrolysis with great ease by this method.

Discussion.

Proton Disintegration.—The probable reaction which results in the production of the 8·4 cm. particles from Li is as pointed out by Cockcroft and Walton (*loc. cit.*)

$$ {}_3Li^7 + {}_1H^1 \longrightarrow {}_2He^4 + {}_2He^4. \qquad (1)$$

If the difference between the initial and the final total masses is the source of the energy of the particles, then it can be shown that the mass of the Li^7 nucleus must be 7·0135, approximately. The mass given by Costa is 7·010, but recent work by Bainbridge* has given the value 7·0146 ± 0·0006 for the atom, corresponding to a nuclear mass of 7·0130 ± 0·0006, in very good agreement with the above assumptions.

The two short ranges which we find when Li is bombarded by protons are much more difficult to explain It was suggested that the 2 cm group found by Cockcroft and Walton (*loc. cit.*) arose as a result of an emission of part of the energy of the disintegration as a γ-ray, and later Traubenberg† claimed to have detected one γ-ray quantum for each disintegration. There are grave difficulties in assigning a mechanism for the production of a γ-ray of the large energy (about 12 × 10⁶ volts) required, and such a penetrating ray would not be detected as easily as ordinary γ-rays. It seems that while the possibility of γ-rays must be borne in mind, the evidence for their existence which we have at present is not conclusive, and it is perhaps more probable that the short range arises in some other manner. For instance, we might imagine the nuclear reaction

$$\left.\begin{array}{ll} & {}_3Li^7 + {}_1H^1 \longrightarrow {}_2He^5 + {}_2He^3 \\ \text{or} & {}_3Li^6 + {}_1H^1 \longrightarrow {}_2He^4 + {}_2He^3 \end{array}\right\} . \qquad (2)$$

Rough estimates of the masses of He^3 find He^5 from Aston's curve, and of the range-energy relation for such particles by analogy with He^4 and H, give energies which are not impossibly different from those obtained. There are other reactions involving neutrons and protons as products of the disintegration which will fill the mass conditions approximately. These possibilities are being explored by searching for γ-rays, for protons, and for neutrons, and by measuring the ratio e/m for the disintegration particles. It is obviously of great importance to study Li^6 and Li^7 separately, and determine which particles arise from each, and since we (*loc. cit.*) have shown that films of lithium less than one atom thick are easily tested there is the possibility of separating sufficient material by mass analysis.

H^2 *Disintegration.*—The absorption curve, fig. 4, shows that the ions of the heavy isotope of hydrogen produce two groups of particles from lithium, the most energetic of which possesses a definite energy corresponding to a mean range of 13·0 cm. This first group probably corresponds with the 14·8 cm. particles reported by Lewis, Livingston and Lawrence. The difference in range may be due to error in one of the measurements, or it may be real and arise from the fact that the energy of the bombarding protons used by them was much greater than ours. The mean range of 13·0 cm. corresponds very nearly to 11·5 million volts, and it is probably produced by the reaction pointed out by Lewis, Livingston and Lawrence

---

* 'Phys. Rev.,' vol. 44, p. 56 (1933).

† 'Z. Physik,' vol. 80, p. 557 (1933).

$$ {}_{3}\mathrm{Li}^{6} + {}_{1}\mathrm{H}^{2} \longrightarrow {}_{2}\mathrm{He}^{4} + {}_{2}\mathrm{He}^{4}. \qquad (3)$$

The two particles must then be expelled in opposite directions with identical energies. The mass of the Li^6 nucleus is 6·0130 according to Bainbridge (*loc. cit.*), and hence in the above reaction a mass disappears of 0·0236 mass units, corresponding to an energy for each particle of $11 \cdot 0 \times 10^6$ volts. The agreement is good. There seems little doubt that this must be the reaction which produces the 13·2 cm. particles.*

The second group of disintegration particles produced from lithium by H^2 ions shows a complete mixture of energies from that corresponding to the shortest observed range of about 1 cm. up to that of the 7·8 cm. particles. We (*loc. cit.*) have reported in a former paper an analysis of the ranges of the particles emitted by boron when bombarded by protons, and in that case also we observed a completely inhomogeneous group of ranges. We were able to show that it was probable that the structure of this group could be satisfactorily explained on the assumption that the B^{11} nucleus broke up into three α-particles when the proton entered, and that these particles escaped in a plane the energy depending on the particular distribution of the particles with angle. By analogy with this boron case we might expect the wide range distribution found for the particles produced from lithium when bombarded by H^2 to be due to the fact that three bodies are emitted in the disintegration. We assume the following reaction

$$ {}_{3}\mathrm{Li}^{7} + {}_{1}\mathrm{H}^{2} \longrightarrow 2\,{}_{2}\mathrm{He}^{4} + {}_{0}n^{1}, \qquad (4)$$

i.e., that the Li^7 nucleus breaks up when an H^2 particle is captured to give two α-particles and a neutron, the total energy released being E. We assume further that the α-particles are ejected with the same energy E_α and at an angle 2θ with respect to one another, fig. 5. The momentum and energy conditions lead to the following expressions for the energy of the neutron and α-particles respectively:

$$\left.\begin{aligned} \mathrm{E}_n &= \mathrm{E}\,.\,\frac{8\cos\theta}{1+8\cos^2\theta} \\ \mathrm{E}_\alpha &= \mathrm{E}/2\,(1+8\cos^2\theta) \end{aligned}\right\}. \qquad (5)$$

---

* This conclusion is made practically certain by the work of Walton and Dee, p. 739.

The reaction (4) leads to the disappearance of 0·0173 mass units if the mass of the Li^7 nucleus is assumed to be 7·0130, as required for the disintegration by protons according to (1). This gives a total energy E available of $16 \cdot 2 \times 10^6$ volts. In fig. 6 we have shown curves calculated from (5) for the variation in the energy of the neutron and α-particle as the angle θ varies from 0° to 90°.

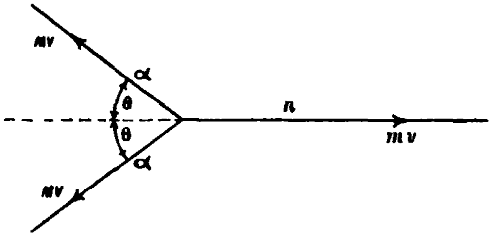


FIG. 5.

It is clear that while the energy of the α-particle varies from 0·056E to 0·5E, the corresponding energy of the neutron varies from 0·889E to zero. Thus we would expect the disintegration products to consist of neutrons with all energies up to about $14 \cdot 4 \times 10^6$ volts, and α-particles with energies between $0 \cdot 9 \times 10^6$ and $8 \cdot 1 \times 10^6$ volts. We have not yet looked for the neutrons, but we have seen that the mixture of α-particles is present and that the maximum energy observed is $8 \cdot 3 \times 10^6$ volts, corresponding to a range of 7·8 cm. The agreement with the calculated value is remarkably good.

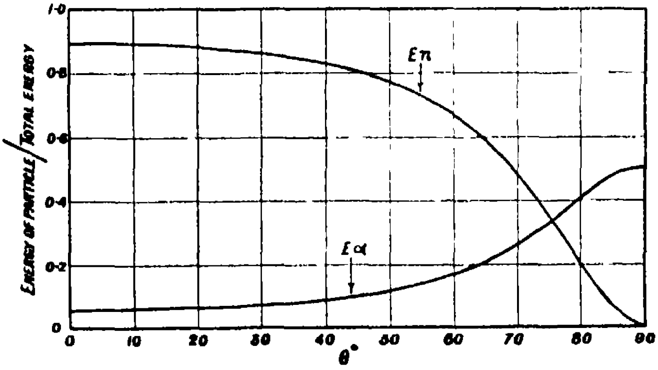


FIG. 6.

Using the curves of fig. 6 it is a simple matter to transform the absorption curve for the group of particles in question, fig. 3, to a corresponding curve in which the number of particles is plotted against the angle θ. The curve so obtained is given in fig. 7. It is clear from this that the maximum of the angular distribution of the α-particles occurs between 70° and 80°, which is the range of angle over which all the disintegration particles possess the same energy, fig. 6. The evidence which we have obtained so far is thus strongly in support of the reaction (4) as the mode of disintegration of Li^7 by bombardment with H^2. We hope to search for neutrons in the immediate future.

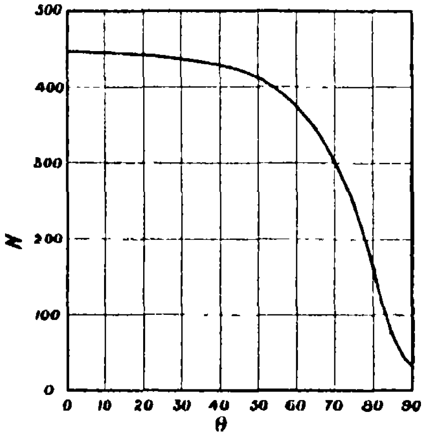


FIG. 7.

We are much indebted to Professor G. N. Lewis for his generous gift of a supply of water containing a large percentage of the heavy isotope of hydrogen —a gift which made some of these experiments possible. We also wish to thank Mr. G. R. Crowe for his valuable technical assistance.

Summary.

(1) The absorption curve for the disintegration particles produced from lithium by protons is shown to consist of three definite parts. A long range of 8·4 ± 0·2 cm. is found, in agreement with Cockcroft and Walton, while there are two short ranges of 1·15 ± 0·2 cm. and 0·65 ± 0·2 cm. mean range respectively, the ratio of the number of particles in the three groups being approximately 1 : 1 : 0·5.

(2) The disintegration of lithium by ions of the heavy isotope of hydrogen was observed. The rate of increase in the number of particles with increase of bombarding energy was much greater than for protons, but in this case no particles were detected with certainty below about 70 kv.

(3) The protons ejected from lithium by ions of heavy hydrogen are shown to fall into two groups, a definite group of 13·0 ± 0·2 cm. mean range, corresponding to an energy of 11·5 million e volts, and a second shorter range group

which is complex and consists of particles of all ranges from about 1·0 cm. up to 7·8 ± 0·2 cm., *i.e.*, with all energies from 1·7 to 8·3 million ε volts.

(4) It is shown that the particles are probably helium nuclei and that the 13·0 cm. group arises from the reaction

$$_3\mathrm{Li}^6 + {}_1\mathrm{H}^2 \longrightarrow {}_2\mathrm{He}^4 + {}_2\mathrm{He}^4,$$

while the complex group is a result of the three-body nuclear reaction

$$_3\mathrm{Li}^7 + {}_1\mathrm{H}^2 \longrightarrow 2\,{}_2\mathrm{He}^4 + {}_0n^1.$$

The masses and energies check within the experimental error.

A Photographic Investigation of the Transmutation of Lithium and Boron by Protons and of Lithium by Ions of the Heavy Isotope of Hydrogen.

By P. I. Dee, M.A., Stokes Student, Pembroke College, and E. T. S. Walton, Ph.D., Clerk Maxwell Scholar, Trinity College, Cambridge.

(Communicated by Lord Rutherford, O.M., F.R.S.—Received August 1, 1933.)

[Plates 14–17.]

1. *Introduction.*

In the preceding paper, p. 722, Oliphant, Kinsey and Rutherford have given an account of the examination of the disintegration of lithium by protons and by ions of the heavy isotope of hydrogen, using electrical counting methods, and have shown that the results of their experiments lend strong support to the views of the modes of disintegration which they there suggest. Certain of these conclusions may be more completely examined by photographing the tracks of the disintegration particles in an expansion chamber and with that object we have made the experiments described below. In the course of this work a great amount of experimental data has been collected which will require a more detailed analysis; in the present paper the photographs described have been selected with the object of testing the above-mentioned theories. It is possible that the photographs show evidence of other modes of disintegration, but in view of the time required for a full analysis, we publish here only an account of the more obvious phenomena.

2. *Experimental Arrangement.*

The apparatus used for the production of the high voltage and its application to the tubes used for accelerating the bombarding particles was that described by Cockcroft and Walton,* and potentials up to about 400 kilovolts were used. The first attempts to work an expansion chamber in conjunction with that apparatus showed that the maximum number of disintegrations produced per second was much too small, and considerable time was spent in attempting to obtain a more intense beam of protons. The form of discharge tube finally adopted was that described by Oliphant and Rutherford.† This has been found much more definite in behaviour than the glass discharge tube used in the early work of Cockcroft and Walton. The total positive ion current measured at the target is at least ten times greater than could be obtained from the latter tube, and the number of disintegrations produced per second has been increased by an even larger factor. It is probable that the new type of tube gives a larger ratio of protons to molecular ions in the beam—this ratio being more liable to fluctuation with the glass discharge tubes. With the present arrangement, using the maximum output of current and voltage, it is possible to obtain over 100 tracks per expansion from a lithium target a few square millimetres in area.

The expansion chamber was of the standard Wilson type 15 cm. in diameter and 5 cm. deep. This type of chamber is not the most suitable for the purpose in hand and it is intended to erect later a completely sealed chamber which can be operated at any pressure. The expansion chamber, cameras, and mechanism requiring attention during use were mounted in a compact manner upon a framework which could be accommodated together with the operator in the lead-lined observation cubicle at the foot of the proton tube. Photographs were taken through the glass roof of the chamber with two cameras mounted together on a board such that the angle made with the vertical by the axis of each camera lens was 30°, the angle between the two optic axes being 20°. The backs of the cameras were tilted relative to the lens axes in the normal manner, the object plane being horizontal and passing through the centre of the chamber. Photographs were taken at a reduction of 1 : 3·5, the lenses being of focal length 3¾ inches. This camera board could be removed as a unit and the plates reinserted in the cameras after development. Measurements were then made upon images of the tracks which were reproduced in space in the relative positions in which they actually occurred.

---

* 'Proc. Roy. Soc.,' A, vol. 136, p. 619 (1932).

† 'Proc. Roy. Soc.,' A, vol. 141, p. 259 (1933).

In these experiments, the proton beam fell upon the target contained in a small tube which passed through the top of the chamber as illustrated in fig. 1. The target was a few square millimetres in area and had a small stopping power. It was inclined to the incoming beam at an angle of 30° and was surrounded by a window system (usually of mica) which had to withstand atmospheric pressure and at the same time admit the products of disintegration into the expansion chamber. Owing to the large proton currents now obtainable it was possible to make these tubes only 1 cm. in diameter and thus very little of the expansion chamber was obscured from the view of the cameras.

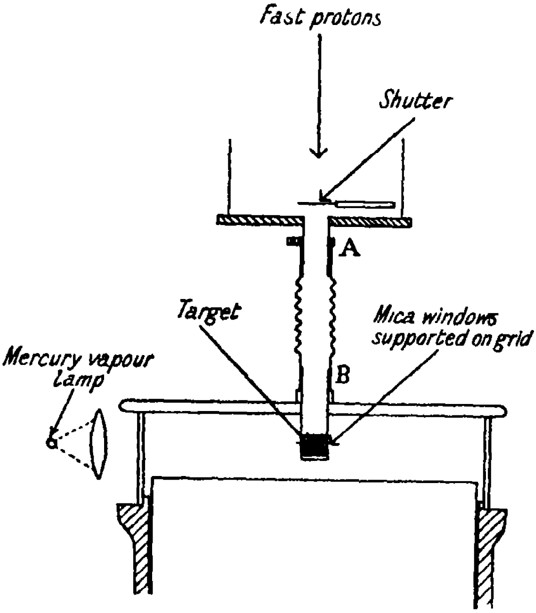


FIG. 1.

To study products of different ranges, grids of various types were inserted in the central hole in the glass roof of the chamber. Connection was made to the end of the main proton tube by a piece of " tombac " tubing held in such compression that although evacuated, its tendency was to expand, and hence any vibration of the chamber during expansion did not impair the vacuum seals. A change of windows or target could be made without admitting air to the whole of the accelerating tubes, by closing a flap in the main proton tube and removing the tombac. The proton beam was shielded from the target by a light shutter which was held closed by an electromagnet inside the proton tube. It was opened at the desired instant relative to the expansion by short circuiting the electromagnet by a thyratron. The grid was charged to the striking voltage through a capacity resistance circuit, variation of the resistance giving the required control over this instant. The striking of the thyratron was also arranged to operate a light relay which produced the illumination, the timing of which was effected by the variation of the resistance in a second circuit which contained some inductance. The use of this electric timing has the advantage of compactness and has been found to be very reproducible in behaviour.

3. *The Disintegration of Lithium by Protons.*

Cockcroft and Walton first showed that the disintegration of lithium under proton bombardment gave rise to particles of 8·4 cm. range,* and to a shorter group of less than 2·0 cm. range.† They showed that the assumed reaction

$$ {}_3\mathrm{Li}^7 + {}_1\mathrm{H}^1 \longrightarrow {}_2\mathrm{He}^4 + {}_2\mathrm{He}^4 \qquad (1)$$

gave good agreement with the 8·4 cm. range, and also showed by scintillation counting that there was evidence of the simultaneous emission of two particles. Kirchner‡ has published photographs obtained in an expansion chamber of the emission of two particles in nearly opposite directions, and has shown that the observed angle between the tracks is in approximate agreement with that calculated from the energy and momenta relations. Kirchner did not arrange for the tracks to end in the expansion chamber, and thus a simultaneous measurement of the ranges of pairs of particles emitted in opposite directions could not be made. In our experiments, four windows of 5·1 cm. stopping power were used mounted upon grid (*a*) of fig. 2. With this arrangement the tracks were a convenient length for measurement.

The constancy of range in this type of disintegration is shown by fig. 3, Plate 14. The difference in range on the two sides of the target is due to the fact that particles emerging on the right have to pass through a piece of mica of 6 mm. stopping power on which was deposited the layer of lithium oxide of 2 mm. stopping power. Figs. 5, 6 and 7, Plate 15, show four disintegrations of this type. In all cases, tests were made as to whether the two tracks passed through a point. The plane containing the two tracks was usually vertical—any departure could be attributed to a scattering from its original vertical direction of the proton responsible for the disintegration.

---

* 'Proc. Roy. Soc.,' A, vol. 137, p. 229 (1932).

† Nature, vol. 131, p. 23 (1933).

‡ 'Sitz. Ber. bayer. Akad. Wiss.,' p. 129 (1933).

The mean value of the sum of lengths of pairs of tracks taken from a number of such photographs was 16·6 cm. The differences in the lengths of each pair in the actual expansion chamber was in agreement with the air equivalent of the path of the right-hand member in the mica used to support the lithium oxide. The average range of the particles emitted in opposite directions was therefore 8·3 cm. which is in good agreement with the value 8·4 cm deduced from the absorption curves, and with the theoretical value of 8·25 cm. obtained by substituting the latest data for the masses of the nuclei concerned in reaction (1). This would seem to furnish complete confirmation of this mode of disintegration. In many experiments tracks appeared without the presence of a related opposite track; this is most probably due to inefficiency of the windows. A further study would be required to be able to affirm that the above reaction occurs in all experiments where the 8·4 cm. particles are produced. Kirchner found no evidence of the short range (less than 2 cm.) group reported by Cockcroft and Walton, this being probably due to the windows used having too great a stopping power. Oliphant, Kinsey and Rutherford, p. 723, have examined this region more fully and report two ranges of about 0·7 and 1·2 cm.

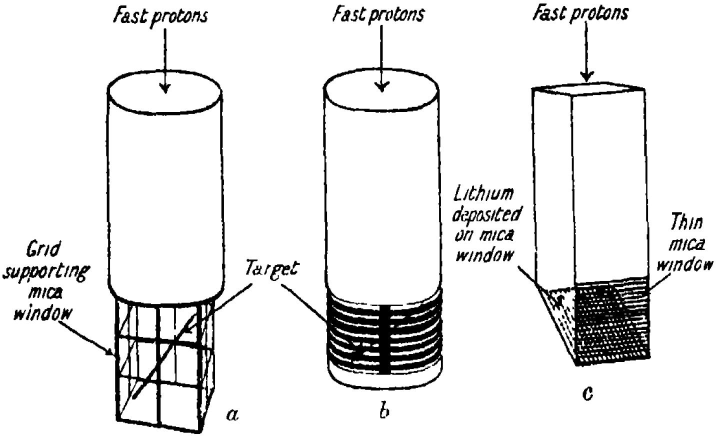


Fig. 2.

In order to photograph such short ranges satisfactorily, it is essential to use windows of less than 4 mm. air equivalent, this necessitating very close spacing of the supporting grid, and consequently a high inefficiency for penetration by incident particles. A mixture of 90% hydrogen and 10% air was used in the chamber in order to increase the lengths of the tracks. An attempt to examine this mode of disintegration was made with the grid shown in fig. 2 (*c*), the slits being 0·3 mm. wide. On the sloping side of the target the particles emitted from the lithium oxide have an unobstructed solid angle for emergence of nearly 2π while particles passing through the opposite vertical face have to pass through the supporting grids. Although these grids had an efficiency of more than 50% for normally incident particles, their finite thickness prevented emergence of particles which fell obliquely upon them. Fig. 8, Plate 15, taken under these conditions, shows clearly the presence of the 12 mm. range and suggests the existence of a group of still shorter range. A number of these longer tracks was measured and gave an average range of 11 mm. The few tracks of still shorter range which were photographed had a length of about 6 mm. In a short run, taken under these conditions, we only obtained one track emerging from the vertical face, and this was the companion to the long range track visible on the other side, both passing obliquely out of the chamber. It is intended to investigate this problem in detail by passing the proton beam into the expansion chamber upon a target suspended in the gas.

4. *The Disintegration of Lithium by Ions of the Heavy Isotope of Hydrogen.*

Lewis, Livingston and Lawrence* have reported the existence of particles of ranges 14·5 cm. and 35·0 cm. when lithium is bombarded by ions of the heavy isotope of hydrogen. Oliphant, Kinsey and Rutherford, p. 727, have described the absorption curve of these particles and have concluded that there is a continuous range distribution up to 7·8 cm. with a homogeneous group of 13·2 cm. range. Lewis, Livingston and Lawrence have suggested the possible reaction

$$_3Li^6 + {}_1H^2 \longrightarrow {}_2He^4 + {}_2He^4 \qquad (2)$$

to account for the 13·2 cm. group. The value calculated from equation (2) using Bainbridge's results is in very close agreement with the observed range.

Lord Rutherford kindly gave us a sample of the heavy isotope of hydrogen which had been presented to him by Professor G. N. Lewis, and this was passed into the discharge tube for the following experiments, the bombarded element being lithium as before. Fig. 4, Plate 14, shows the general type of events obtained under these conditions, the windows having the same stopping power as for fig. 3, Plate 14, which was for proton bombardment. The existence of a group with all ranges up to about 8 cm. is clearly shown, together with particles which passed out to the walls of the chamber and had therefore a range of more than 10 cm.

---

* 'Phys. Rev.,' vol. 44, p. 55 (1933).

Figs. 9, 10 and 11, Plate 16, show examples with smaller ion currents and strongly suggest the emission of pairs of long range particles in opposite directions as is required by the conservation of momentum in reaction (2). About 100 similar photographs have been taken and the number of such opposite pairs is far greater than could be attributed to chance. Thicker windows of 10·4 cm. stopping power were then used on a more efficient grid system, and figs. 12 and 13, Plate 17, show a pair of stereoscopic photographs of a disintegration obtained under these conditions. The tracks then ended in the gas and permitted measurement of ranges. The reduced range of each of this particular pair of particles was 13·4 cm., in good agreement with the value of 13·2 obtained by Oliphant, Kinsey and Rutherford, and with the value obtained from equation (2). This particular mode of disintegration would thus seem to be definitely confirmed. Fig. 14, Plate 17, shows another pair of particles of this range, together with a track which passes out to the wall of the chamber and has therefore a range of at least 16 cm. Many other similar tracks were obtained, and the smaller ionization along them would suggest that they may constitute the 35 cm. group reported by Lawrence, Livingston and Lewis (*loc. cit.*). In a brief examination of these tracks, no obvious correlation has been noticed between their direction of emission and those of other particles. Among the tracks of about 8 cm. range, we have observed a few forks which were measured, and gave approximate values of the ratio of the masses of the colliding particles in agreement with the ratios He to O or He to N.

5. *Disintegration of Boron by Proton Bombardment.*

Cockcroft and Walton showed that boron under proton bombardment emitted particles of ranges up to about 5 cm. and suggested that three α-particles were produced in the disintegration according to the equation

$$ {}_5\mathrm{B}^{11} + {}_1\mathrm{H}^1 \longrightarrow {}_2\mathrm{He}^4 + {}_2\mathrm{He}^4 + {}_2\mathrm{He}^4 \tag{3}$$

Oliphant and Rutherford* showed that a careful examination of the absorption curve gave strong support to this hypothesis and Kirchner (*loc. cit.*) has published a photograph of three particles emitted at approximately 120°, but does not state the ranges of the particles or whether the three tracks show conservation of momentum.

---

* 'Proc. Roy. Soc.,' A, vol. 141, p. 259 (1933).

Using a thin piece of pyrex glass (which contains boron) of 2·5 mm. stopping power as target and a mica window of 6·4 mm. stopping power mounted on the grid shown in fig. 2 (*b*), we have taken over 100 photographs. In order to make the average track a suitable length, a mixture of 80% hydrogen and 20% air was used in the chamber. The slots of this grid were 0·4 mm. wide and the efficiency of the grid for normally incident particles was about 70% Fig. 15, Plate 17, is typical of this run. On these plates there were many cases of groups of three particles which appeared to be due to simultaneous emission of three particles, but after careful measurement on the tracks reproduced in space it was found that although there were cases of three tracks lying in the same plane and passing through a point, the majority of these did not satisfy the momentum relations as accurately as could be reasonably expected. The possibility of slight deflections in the mica windows of particles of such small residual range must not be overlooked. For the study of a three-body emission, it is essential to reduce the inefficiency of the windows to a further extent and a full discussion of this disintegration is therefore postponed to a later date.

This research was carried out with the high voltage installation set up by Cockcroft and Walton and we are much indebted to Dr. Cockcroft for use of this apparatus and for valuable suggestions. We wish to express our gratitude to Lord Rutherford for his interest in the work, and for the benefit we have derived from many helpful discussions. One of us (E. T. S. W.) wishes to acknowledge a grant from the Department of Scientific and Industrial Research. Our thanks are also due to Mr. W. Birtwhistle for much technical assistance.

Summary.

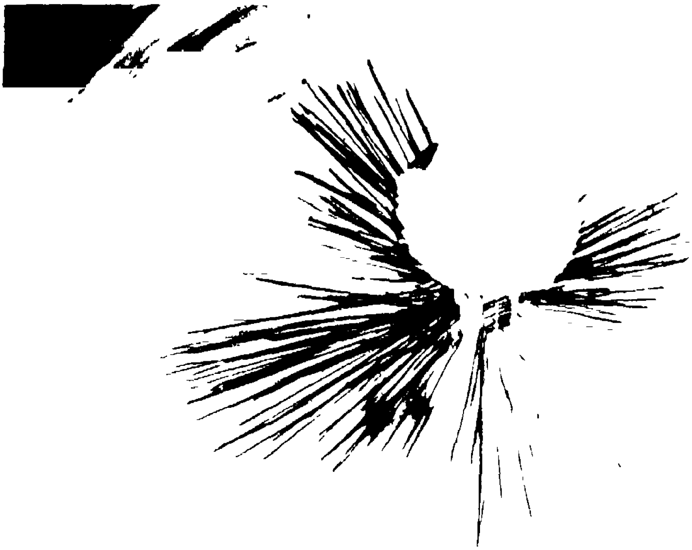
Photographs have been taken, using an expansion chamber and double camera, of the tracks emitted in the transmutations of lithium and boron by protons and of lithium by ions of the heavy isotope of hydrogen. By measurement of the tracks reproduced in space from the photographs taken, the modes of disintegration

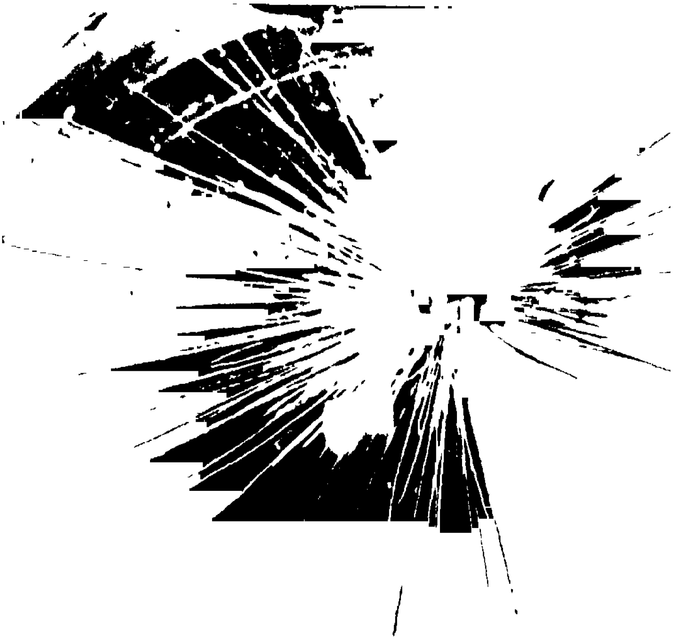
$${}_3\mathrm{Li}^7 + {}_1\mathrm{H}^1 \longrightarrow {}_2\mathrm{He}^4 + {}_2\mathrm{He}^4$$

and

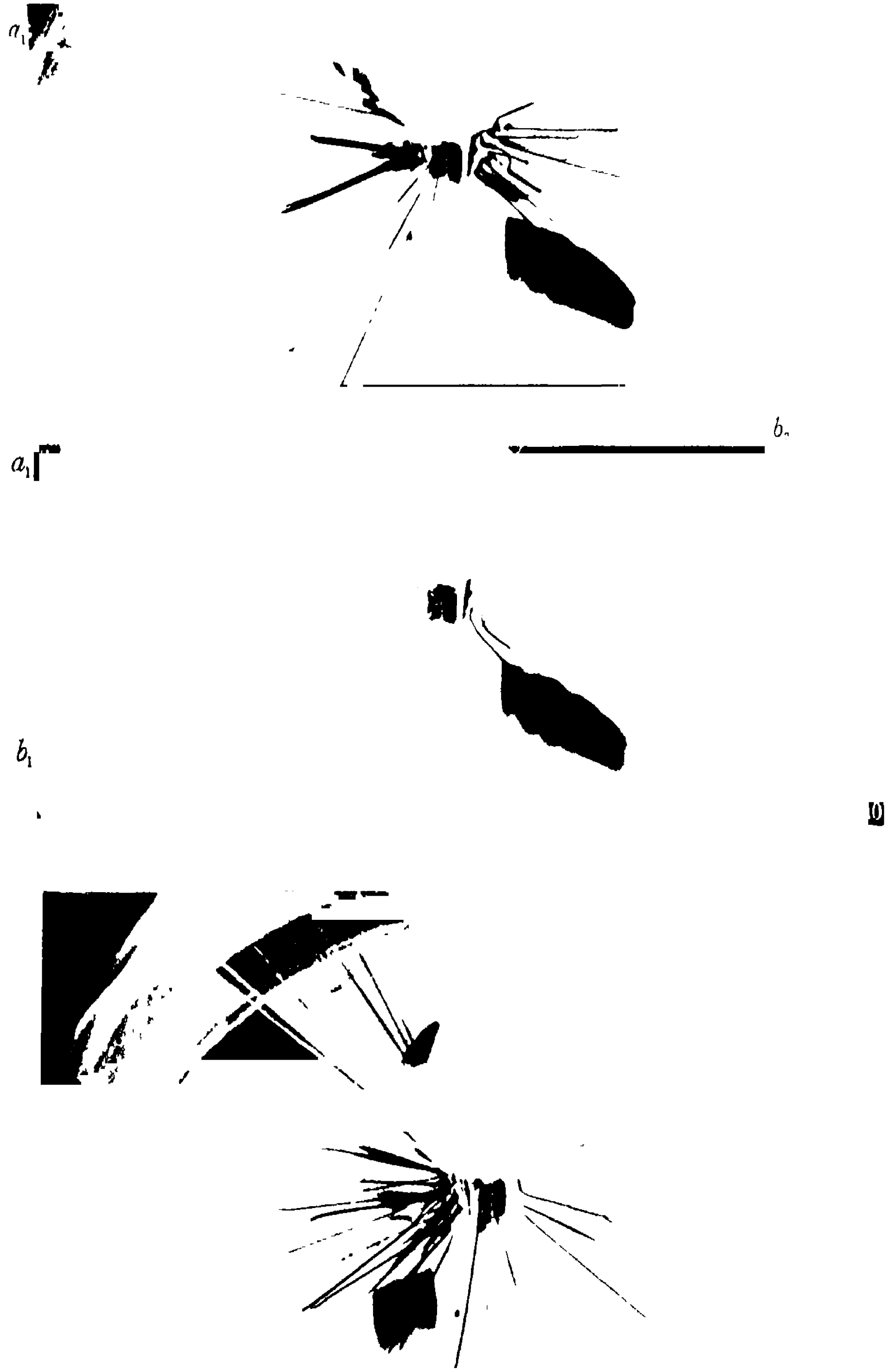
$${}_3\mathrm{Li}^6 + {}_2\mathrm{H}^2 \longrightarrow {}_2\mathrm{He}^4 + {}_2\mathrm{He}^4$$

have been definitely confirmed.

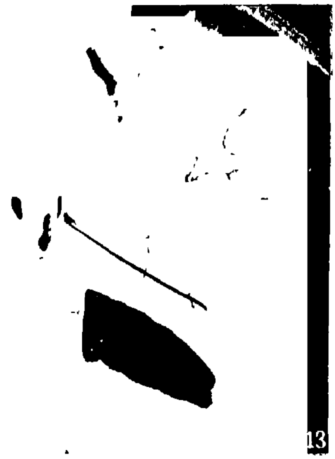




a_2'
b_1
5
d_1
c_1
c_2
d_2


a_1
a_1
b_2
b_1
0

12



13

14

15

16

DESCRIPTION OF PLATES.

At the centre of each photograph the window system surrounding the target can be seen, and above this the tube down which passes the stream of fast ions. The emergent tracks are produced by the particles emitted in the disintegrations. Except where otherwise stated the stopping power of the windows used was 5.1 cm. of air.

Figs. 3–14 are reproduced three-quarter natural size, figs. 15 and 16 at half natural size. Tracks are referred to by letters at the margin in their direction from the target. The white patch on each photograph is the image in the floor of the chamber of the target tube.

PLATE 14.

FIG. 3.—Protons incident on thin lithium target, the α-particles emitted showing a constant range of 8·4 cm. The particular fan shape is due to oblique passage of the α-particles through the mica windows and interception by the supporting grid. The particles on the right have shorter length in the chamber as they have to pass through the mica which carries the lithium oxide.

FIG. 4.—Ions of the heavy isotope of hydrogen incident upon thin lithium target. This photograph was taken under the same conditions as fig. 3 and shows the existence of particles which pass out to the walls of the chamber and have therefore a range greater than 10 cm. The continuous range of less than 8 cm is also obvious. Some of the thinner tracks may be due to fast protons, but the density of the track under these conditions of photography is not a certain guide as to the nature of the ionizing particle.

PLATE 15

FIGS. 5 and 6.—A stereoscopic pair showing the disintegration of lithium by proton bombardment with the emission of pairs of α-particles in opposite directions. Tracks a_1, a_2 belong to one disintegration, b_1, b_2 to another. These tracks occupy the positions a'_1, a'_2 and b'_1, b'_2 in the other photograph, illustrating the manner in which the pair of photographs taken from different positions shows that the tracks are colinear and not merely coplanar.

FIG. 7.—Another photograph showing two separate disintegrations of lithium by protons. The tracks c_1, c_2, d_1, d_2 being the opposite pairs

FIG. 8.—The short range group from lithium under proton bombardment photographed in a mixture 90% hydrogen and 10% air, the particles passing through mica windows of 4 mm. stopping power. An opposite pair of long range particles is also present which pass obliquely out of the illumination. The length of the short range tracks reduced to standard air is 11 mm.

PLATE 16.

FIG. 9.—Disintegration of lithium by ions of the heavy isotope of hydrogen, showing the emission of two particles (a_1, a_2) in opposite directions passing out of the expansion chamber and with ranges therefore greater than 10 cm. ($Li^6 + H^2 \rightarrow 2He^4$). The thin long track ($b$) is probably a fast proton.

FIG. 10.—This photograph shows the type of disintegration described above, fig. 9, and also a disintegration with the emission of a pair of opposite 8·4 cm. particles (b_1, b_2), probably due to the presence of protons in the positive ion beam.

FIG. 11.—This shows several cases of particles with range greater than 10 cm. lying in opposite directions. There are also particles with different ranges less than 8 cm. ending in the chamber.

PLATE 17.

FIGS. 12 and 13.—This pair of photographs was taken with ions of the heavy isotope of hydrogen incident upon thin lithium, the windows having a stopping power of 10 cm. which made the particles end in the chamber. The ranges of this opposite pair are 13 4 cm. corresponding to the transmutation of $Li^6 + H^2$ into two helium atoms.

FIG. 14.—This was taken under the same conditions as the preceding but showing a fine track, probably due to a proton, passing out of the chamber and with range therefore greater than 16 cm. This photograph also shows another pair of particles of 13 cm. range emitted in opposite directions.

FIG. 15.—This is a typical photograph of the tracks produced by the bombardment of boron by protons. Three of the tracks lie in a plane and pass through a point, but measurement shows that conservation of momentum does not hold and hence the tracks probably come from different disintegrations.

FIG. 16.—This shows a pair of oppositely directed particles from boron under proton bombardment and may illustrate one of the cases discussed by Rutherford and Oliphant where two particles come out in nearly opposite directions and the third receives very little energy.

FIGS. 15 and 16.—These were taken in a mixture of 80% hydrogen and 20% air; stopping power of windows = 6.4 mm.

INDEX TO VOL. CXLI. (A.)

Absorption coefficient of a partially degenerate gas (Swirles), 554.
Acetaldehyde, thermal decomposition (Fletcher and Hinshelwood), 41.
Alpha-particles, collisions with fluorine nuclei (Feather), 194.
Aluminium films, structure (Finch and Quarrell), 398.
Appleton (E. V.) and others. Discussion on the Ionosphere, 697.
Arsonic acid, possessing molecular dissymmetry (Gibson and Levin), 494.
Atomic wave functions, calculations (Hartree), 282.
Awbery (J. H.) and Griffiths (E.) The Heats of Combustion of Carbon Monoxide in Oxygen and of Nitrous Oxide in Carbon Monoxide at Constant Pressure, 1.

Banerjee (K.) Determinations of the Signs of the Fourier Terms in Complete Crystal Structure Analysis, 188.
Barnett (L.) *See* Gibson and Barnett.
Bartlett (M. S.) Probability and Chance in the Theory of Statistics, 518.
Bell (J.) The Emission of Electrons from Tungsten and Molybdenum under the Action of Soft X-rays from Copper, 641.
Blair (G. W. Scott) *See* Schofield and Scott Blair.
Boron, photographic investigation of transmutation by protons and ions (Dee and Walton), 733.
Brata (L.) Emission of Metallic Ions from Oxide Surfaces, I, 454.
Brata (L.) *See also* Powell and Brata.
Broadway (L. F.) Experiments on Molecular Scattering in Gases. II—The Collision of Sodium and Potassium Atoms with Mercury, 634.
Broadway (L. F.) *See also* Fraser and Broadway.
Bullard (E. C.) The Observation of Gravity by Means of Invariable Pendulums, 233
Bunn (C. W.) Adsorption, Orientated Overgrowth and Mixed Crystal Formation, 567.
Butane-air mixtures, spontaneous ignition (Townend and Mandlekar), 484.
Butler (J. A. V.) and Thomson (D. W.) The Behaviour of Electrolytes in Mixed Solvents, V, 86.

Calculus of variations, approximation by polygons (Young), 325.
Carbon dioxide, bomb calorimeter determination of heat of formation (Fenning and Cotton), 17.
Carbon monoxide, heat of combustion in oxygen (Awbery and Griffiths), 1.
Chain reaction between hydrogen and oxygen, upper pressure limit (Grant and Hinshelwood), 29.
Chaudhri (R. M.) Ionization of Mercury Vapour by Positive Ions of Mercury and Potassium, 386.
Childs (E. C.) and Massey (H. S. W.) The Scattering of Electrons by Metal Vapours, I, 473.
Conduction of heat in powders (Kannuluik and Martin), 144.
Conductivities, thermal and electrical, of metals (Kannuluik), 159.
Conductors and insulators, electronic properties (Fowler), 56.
Cotton (F. T.) *See* Fenning and Cotton.

Crystal formation, adsorption and overgrowth in mixed (Bunn), 567.

Crystal structure analysis, signs of Fourier terms (Banerjee), 188.

Crystal structure of durene (Robertson), 594.

Curtis (W. E.) and Evans (S. F.) The Spectra of the Halogen Molecules. I—Iodine, 603.

Dee (P. I.) and Walton (E. T. S.) A Photographic Investigation of the Transmutation of Lithium and Boron by Protons and of Lithium by Ions of the Heavy Isotope of Hydrogen, 733.

Ditchburn (R. W.) The Deposition of Sputtered Films, 169.

Durene, X-ray analysis of the crystal structure (Robertson), 594.

Electrolytes, behaviour in mixed solvents (Butler and Thomson), 86.

Electromagnetic fields due to variable electric charges (Fock), 550.

Electrons, emission from tungsten and molybdenum (Bell), 641.

Electrons, scattering by metal vapours (Childs and Massey), 473.

Ellis (C. D.) and Mott (N. F.) Energy relation in the β-ray Type of Radioactive Disintegration, 502.

Epidemics, mathematical theory and study of endemicity (Kermack and McKendrick), 94.

Evans (S. F.) *See* Curtis and Evans.

Feather (N.) Collisions of α-particles with Fluorine Nuclei, 194.

Fenning (R. W.) and Cotton (F. T.) A Bomb Calorimeter Determination of the Heats of Formation of Nitrous Oxide and Carbon Dioxide, 17.

Finch (G. I.), Munson (C. A.), Stuart (N.) and Thomson (G. P.) The Catalytic Properties and Structure of Metal Films, I, 414.

Finch (G. I.) and Quarrell (A. G.) The Structure of Magnesium, Zinc and Aluminium Films, 398.

Films, structure of metal (Finch and Quarrell), 398, and (Finch and others), 414.

Films, sputtered, deposition (Ditchburn), 169.

Fletcher (C. J. M.) and Hinshelwood (C. N.) The Thermal Decomposition of Acetaldehyde and the Existence of Different Activated States, 41.

Flint (H. T.) A New Presentation and Interpretation of the Quantum Equations, 363.

Flow past circular cylinders at low speeds (Thom), 651.

Fock (V.) On the Electromagnetic Fields due to Variable Electric Charges and the Intensities of Spectrum Lines according to the Quantum Theory, 550.

Fowler (R. H.) Notes on Some Electronic Properties of Conductors and Insulators, 56.

Fraser (R. J. G.) and Broadway (L. F.) Experiments on Molecular Scattering in Gases. I—The Method of Crossed Molecular Beams, 626.

Gibson (C. S.) and Barnett (L.) An Optically Active Arsonic Acid Possessing Molecular Dissymmetry. Resolution of *dl-spirobis*-3 : 5-Dioxan-4 : 4′-di (phenyl-p-arsonic acid), 494.

Goldsbrough (G. R.) Ocean Currents Produced by Evaporation and Precipitation, 512.

Grant (G. H.) and Hinshelwood (C. N.) The Upper Pressure Limit in the Chain Reaction between Hydrogen and Oxygen, 29.

Gravity observations by invariable pendulums (Bullard), 233.

Griffiths (E.) *See* Awbery and Griffiths.

Gurney (R. W.) Internal Photoelectric Absorption in Halide Crystals, 209.

Hallimond (A. F.) and Herroun (E. F.) Laboratory Determinations of the Magnetic Properties of Certain Igneous Rocks, 302.

Hartree (D. R.) Results of Calculations of Atomic Wave Functions. I—Survey and Self-consistent Fields for Cl^- and Cu^+, 282.

Harper (W. R.) On the Ionization of Light Gases. I and II, 669, 686.

Heat conduction in powders (Kannuluik) and Martin, 144.

Herroun (E. F.) *See* Hallimond and Herroun.

Hinshelwood (C. N.) *See* Fletcher and Hinshelwood.

Hinshelwood (C. N.) *See also* Grant and Hinshelwood.

Ignition, spontaneous, of inflammable gas-air mixtures (Townend and Mandlekar), 484.

Igneous rocks, magnetic properties (Hallimond and Herroun), 302.

Iodine, adsorption by potassium iodide (Whipp), 217.

Iodine, spectrum (Curtis and Evans), 603.

Ionization of Light Gases. I and II (Harper), 669, 686.

Ionosphere, discussion (Appleton and others), 697.

Ions, metallic, emission from oxide surfaces (Brata, and Powell and Brata), 454, 463.

Kannuluik (W. G.) and Martin (L. H.) Conduction of Heat in Powders, 144.

Kannuluik (W. G.) The Thermal and Electrical Conductivities of Several Metals between −183° C and 100° C. Appendix by C. E. Eddy and T. H. Oddie, 159

Kaye (G. W. C.) and Sherratt (G. G.) The Velocity of Sound in Gases in Tubes, 123.

Kermack (W. O.) and McKendrick (A. G.) Contributions to the Mathematical Theory of Epidemics, III, 94.

Kinsey (B. B.) *See* Oliphant, Kinsey and Rutherford.

Lithium, photographic investigation of transmutation by protons and ions (Dee and Walton), 733.

Lithium, transmutation by protons and ions (Oliphant, Kinsey and Rutherford), 722.

McCrea (W. H.) and Newing (R. A.) Boundary Conditions for the Wave Equation, 216.

McKendrick (A. G.) *See* Kermack and McKendrick

Magnesium films, structure (Finch and Quarrell), 398.

Mandlekar (M. R.) *See* Townend and Mandlekar

Martin (L. H.) *See* Kannuluik and Martin.

Massey (H. S. W.) and Mohr (C. B. O.) Free Paths and Transport Phenomena in Gases and the Quantum Theory of Collisions, I, 434.

Massey (H. S. W.) *See also* Childs and Massey.

Mercury vapour, ionization by positive ions (Chaudhri), 386.

Metal films, catalytic properties (Finch and others), 414.

Metallic ions, emission from oxide surfaces (Brata, and Powell and Brata), 454, 463.

Mohr (C. B. O.) *See* Massey and Mohr.

Molecular Scattering in Gases, I and II (Fraser and Broadway, and Broadway), 626, 634.

Molybdenum, emission of electrons from, under the action of soft X-rays from copper (Bell), 641.

Mott (N. F.) *See* Ellis and Mott.

Murison (C. A.) *See* Finch and others.

Neuber (H.) New Method of Deriving Stresses Graphically from Photo-Elastic Observations, 314.

Newing (R. A.) *See* McCrea and Newing.
Nitrogen nucleus, artificial disintegration (Pollard), 375.
Nitrogen tetroxide infra-red absorption spectrum and structure of the molecule (Sutherland), 342.
Nitrous oxide, bomb calorimeter determination of heat of formation (Fenning and Cotton), 17.
Nitrous oxide, heat of combustion in carbon monoxide (Awbery and Griffiths), 1.

Ocean currents produced by evaporation and precipitation (Goldsbrough), 512.
Oliphant (M. L. E.), Kinsey (B. B.) and Rutherford (Lord). The Transmutation of Lithium by Protons and Ions of the Heavy Isotope of Hydrogen, 722.
Oliphant (M. L. E.) and Rutherford (Lord) Experiments on the Transmutation of Elements by Protons, 259.

Photoelectric absorption in halide crystals, internal (Gurney), 209.
Pollard (E. C.) Experiments on the Protons produced in the Artificial Disintegration of the Nitrogen Nucleus, 375.
Powell (C. F.) and Brata (L.) Emission of Metallic Ions from Oxide Surfaces, II, 463.

Quantum equations, new presentation and interpretation (Flint), 363.
Quantum theory of collisions (Massey and Mohr), 434.
Quantum theory, intensities of spectrum lines (Fock), 550.
Quarrell (A. G.) *See* Finch and Quarrell.

Radioactive disintegration, energy relation in β-ray type (Ellis and Mott), 502.
Raman effect at very low temperatures (Sutherland), 535.
Robertson (J. M.) X-ray analysis of the Crystal Structure of Durene, 594.
Rutherford (Lord). *See* Oliphant, Kinsey and Rutherford.
Rutherford (Lord). *See also* Oliphant and Rutherford.

Scattering of electrons by metal vapours (Childs and Massey), 473.
Scattering, molecular, in gases (Fraser and Broadway, and Broadway), 626, 634.
Schofield (R. K.) and Scott Blair (G. W.) The Relationship between Viscosity, Elasticity and Plastic Strength of a Soft Material as Illustrated by some Mechanical Properties of Flour Dough, III, 72.
Sherratt (G. G.) *See* Kaye and Sherratt.
Sound, velocity in gases in tubes (Kaye and Sherratt), 123.
Spectra of halogen molecules (Curtis and Evans), 603.
Spectrum of nitrogen tetroxide, infra-red absorption (Sutherland), 342.
Statistics, theory (Bartlett), 518.
Stresses derived graphically from photo-elastic observations (Neuber), 314.
Stuart (N.) *See* Finch and others.
Sutherland (G. B. B. M.) *Experiments on the Raman Effect at very Low Temperatures*, 535.
Sutherland (G. B. B. M.) The Infra-Red Absorption Spectrum of Nitrogen Tetroxide and the Structure of the Molecule, 342.
Swirles (B.) The Coefficients of Absorption and Opacity of a Partially Degenerate Gas, 554.

Thermal and electrical conductivities of several metals (Kannuluik), 159.

Thom (A.) The Flow Past Circular Cylinders at Low Speeds, 651.
Thomson (D. W.) *See* Butler and Thomson.
Thomson (G. P.) *See* Finch and others.
Townend (D. T. A.) and Mandlekar (M. R.) The Influence of Pressure on the Spontaneous Ignition of Inflammable Gas-Air Mixtures, I, 484.
Transmutation of elements by protons (Oliphant and Rutherford), 259, and 722.
Transmutation of lithium and boron, photographic investigation (Dee and Walton), 733.
Transport phenomena and free paths in gases (Massey and Mohr), 434.
Tungsten, emission of electrons from, under the action of soft X-rays from copper (Bell), 641.

Velocity of sound in gases in tubes (Kaye and Sherratt), 123.
Viscosity, elasticity and plastic strength of a soft material (Schofield and Scott Blair), 72.

Walton (E. T. C.) *See* Dee and Walton.
Wave equation, boundary conditions (McCrea and Newing), 216.
Whipp (B.) The Adsorption of Iodine by Potassium Iodide, 217.

Young (L. C.) On Approximation by Polygons in the Calculus of Variations, 325.

Zinc films, structure (Finch and Quarrell), 398.

END OF THE ONE HUNDRED AND FORTY-FIRST VOLUME (SERIES A)